A STUDY OF THE TECHNOLOGICAL, INSTRUCTIONAL, AND MOTIVATIONAL

FACTORS AFFECTING PHR CERTIFICATION EXAM OUTCOMES

David M. Bonner, B.A., M.S.A.

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2012

APPROVED:

Jerry Wircenski, Major Professor and
    Program Coordinator
Donna E. Ledgerwood, Minor Professor
Jeff Allen, Committee Member
Kim Nimon, Committee Member
Michael Spector, Chair of the Department of
    Learning Technologies
Linda Schamber, Acting Dean of the College
    of Information
James D. Meernik, Acting Dean of the
    Toulouse Graduate School

Bonner, David M. <u>A study of the technological, instructional, and motivational factors affecting PHR certification exam outcomes</u>. Doctor of Philosophy (Applied Technology and Performance Improvement), May 2012, 101 pp., 10 tables, 5 illustrations, references, 96 titles.

Although previous studies have considered the factors affecting other certification exam outcomes, they have not examined those that are related to performance on the Professional in Human Resources (PHR) exam. In response to that need, this study specifically investigates technology and training factors that affect self-efficacy and self-set goals, and through them, influence PHR certification exam results. The target population for the study consisted of recent examinees who had taken a formal PHR examination preparation class or used another form of exam preparation training. The survey results were analyzed using partial least squares modeling techniques, and mediation effects were then tested. The results demonstrated that PHR training self-efficacy affected PHR exam self-efficacy and self-set goals. These factors then had an impact on PHR exam scores. Also, the results of task-technology fit were indirectly related to PHR training self-efficacy through a multiple mediation model that included the instructional factor of time on task and the technology factor of perceived usefulness. Surprisingly, time spent on practice exam questions was found to be negatively related to PHR certification exam scores. Finally, instructional feedback indirectly affected outcomes through its positive relationship to self-set goals. The results of the research should help training professionals and examinees in structuring PHR exam training and preparation activities. They also suggest avenues for improving outcomes in other similar types of training.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1

INTRODUCTION

Background

In human resource management, as in other professional fields, there is a trend toward increased certification (Aguinis, Michaelis, & Jones, 2005). Certification programs play an important role in employment decisions because they "signal to potential employers that an individual has mastered a specific body of knowledge" (p. 160). The Professional in Human Resources or PHR® certification (Human Resource Certification Institute, Alexandria, VA, www.hrci.org) is an example of such a program. PHR exam preparation courseware and materials available from the Society for Human Resource Management or SHRM® (Society for Human Resource Management, Alexandria, VA, www.shrm.org) have been used by hundreds of companies and universities as well as thousands of HR professionals and students to prepare for the PHR exam and develop competencies in the HR profession (Forman & Cohen, 1999).

The body of knowledge underpinning the PHR exam is maintained and administered by the Human Resource Certification Institute or HRCI® (Human Resource Certification Institute, Alexandria, VA, www.hrci.org). The HRCI (2008) administers five credentials and accompanying certification programs. These include the Professional in Human Resources (PHR®), Senior Professional in Human Resources (SPHR®), Global Professional in Human Resources (GPHR®), and two state-specific certifications for California (PHR-CA® and SPHR-CA®) (Human Resource Certification Institute, Alexandria, VA, www.hrci.org). Forman and Cohen (1999) described the SHRM Learning System® (Society for Human Resource Management, Alexandria, VA, www.shrm.org), and by inference, the HRCI body of knowledge, in terms of a knowledge base with a specific focus, not designed to include everything there is to

know about HR theory or practice. They also pointed out that it does not include application or skills-based information, nor does it provide information on how to implement these knowledge-base practices within a workplace setting. Instead, Wiley (1999) described the HRCI programs as having the goal of establishing inputs of knowledge, skills, and attitudes in HR professionals who in turn bring value to their organizations. The purpose of a PHR, SPHR, or GPHR certification is to "show that the holder has demonstrated mastery of the domestic or international HR knowledge base and, through recertification, has accepted the challenge to stay informed of new developments in the HR field" (HRCI, 2011b, p.2).

Need for the Study

As of January 2011, more than 119,000 HR professionals had earned one of the five credentials that HRCI administers (HRCI, 2011a). That includes more than 67,000 who had earned the PHR certification. As noted earlier, however, the PHR exam itself is underpinned by a knowledge base with a specific focus. This leads to several questions, such as how an examinee should prepare for the exam and what training professionals could do to design better PHR certification training programs. Unfortunately, no study could be found which examined the characteristics or the training activities that correlated with success on the PHR exam. The research that was located focused more on the relationships between the HRCI body of knowledge and either workplace outcomes (Rynes, Colbert, & Brown, 2002) or college curriculum development (Sincoff & Owen, 2004). One recent study by Fertig (2011) investigated the motivational aspects of PHR certification, but not the exam preparation activities. Similarly, a study by Fertig, Zeitz, and Blau (2009) considered motivation to obtain a competency certification, using the PHR exam as an example, but it did not address exam preparation.

Studies of other certification programs have been done, however, that looked at examinee ability and preparation characteristics and their correlation with exam outcomes. Several of these studies demonstrated the effects of two specific training factors in a certification program setting. The first factor was suggested by several studies (Grange, Hampton, Cutler, Langdon, & Ryan, 2003; Grant, Ciccotello, & Dickie, 2002), which found that testing outcomes varied based on training approach. The second factor addressed in studies of certification exams was the impact of practice exams (Harrington, Davis, & Harrington, 1992; Simonsson, Poelzer, & Zeng, 2000). These studies showed that the use of practice exams was positively correlated with certification test results.

While these results are interesting, and the second one makes sense intuitively, they still provide little information about what can be done to improve the probability of success on the PHR exam. According to HRCI (2011a) statistics, the probability of success has ranged from 56% to 60% for PHR exams taken during the period from May 2008 through January 2011. As a result, it is proposed that more research is needed to add to the body of knowledge, specifically investigating the training factors that have an impact on PHR certification exam outcomes and the mechanisms at work. The results of the research can help training professionals and examinees in structuring PHR exam preparation activities. They may also suggest ways in which training for other competency testing programs could be improved.

<center>Theoretical Framework</center>

Overall Theoretical Framework

The overall theoretical framework for this study is illustrated in Figure 1. The framework is based on a model (see Figure 2) that describes a set of relationships among goals, self-efficacy and performance (Latham & Locke, 1991; Locke, Frederick, Lee, & Bobko, 1984). Locke

<center>3</center>

(1991) referred to this set of relationships as the motivation hub. The antecedents of self-efficacy, and self-set goals were selected by referencing prior studies that addressed precursors to motivation and psychological learning processes. Colquitt, LePine, and Noe (2000) suggested that both individual and situational characteristics can act as antecedents of motivation and learning, and their meta-analysis identified a number of factors in each category. Because of its intended focus on the subject of interventions, this study has incorporated a framework suggested by Alavi and Leidner (2001) for technology-mediated learning research (see Figure 3). The researchers defined the scope of technology mediation broadly to include access to learning materials, peer interactions, and/or instructor interactions. Alavi and Leidner proposed a model with two categories of antecedents, instructional strategies and information technology, which affect psychological learning processes, such as motivation, that in turn predict learning outcomes. Their framework suggested that these categories exist within a learning context.

Next, three of Chickering and Gamson's (1987) seven instructional best practices, time on task, feedback, and expectations, were selected as the instructional strategies because of both empirical results and theoretical support for their relationship to self-efficacy or self-set goals. These three factors were added to the two suggested by prior certification exam research, type of training, and time spent on practice exam questions. Finally, task-technology fit (Goodhue & Thompson, 1995) and perceived usefulness (Davis, 1989) were selected to measure the effect of information technology in a training environment. Their selection was based on both empirical and theoretical support for their roles in affecting performance and utilization of the technology.
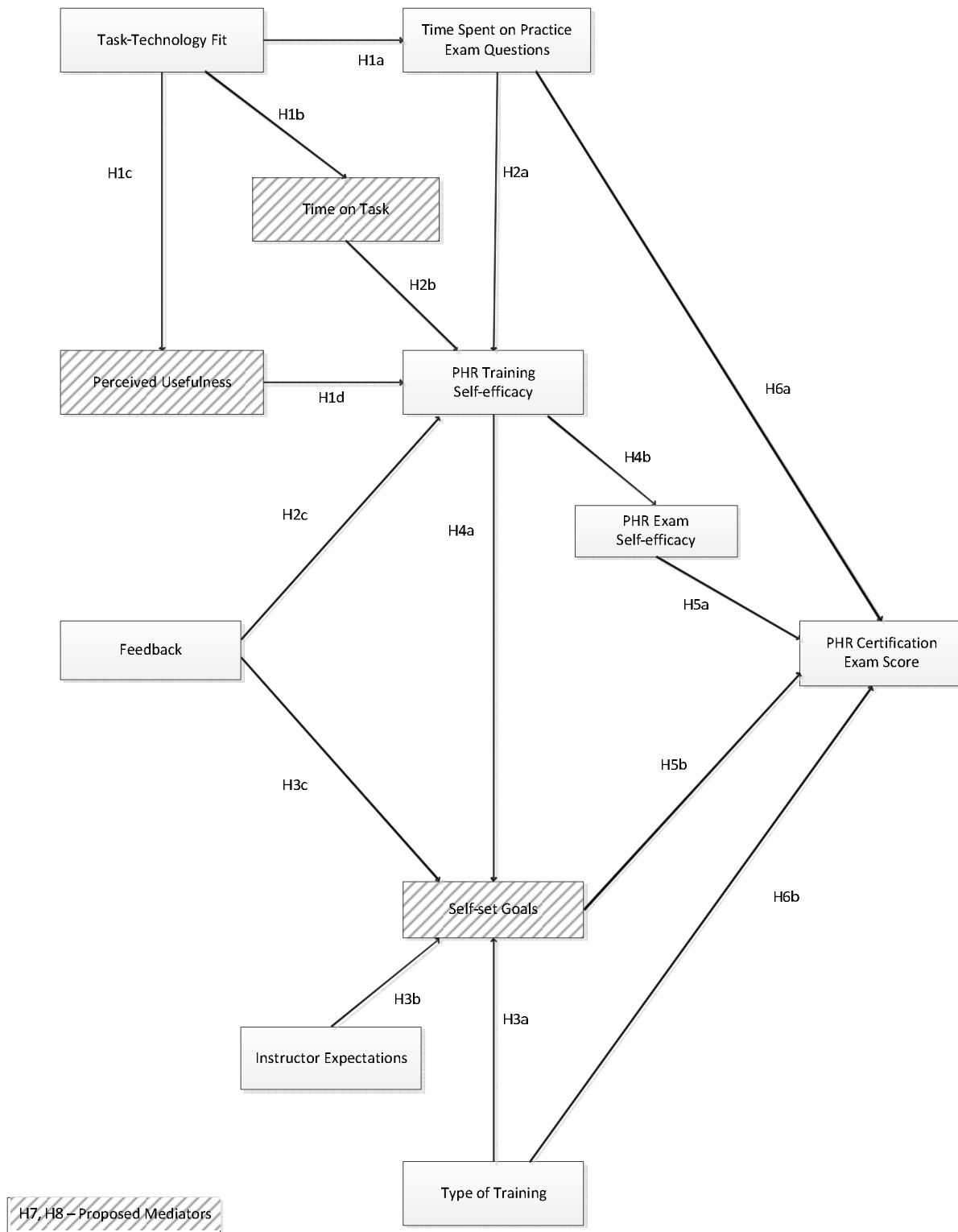
Figure 1. Hypothesized relationships.
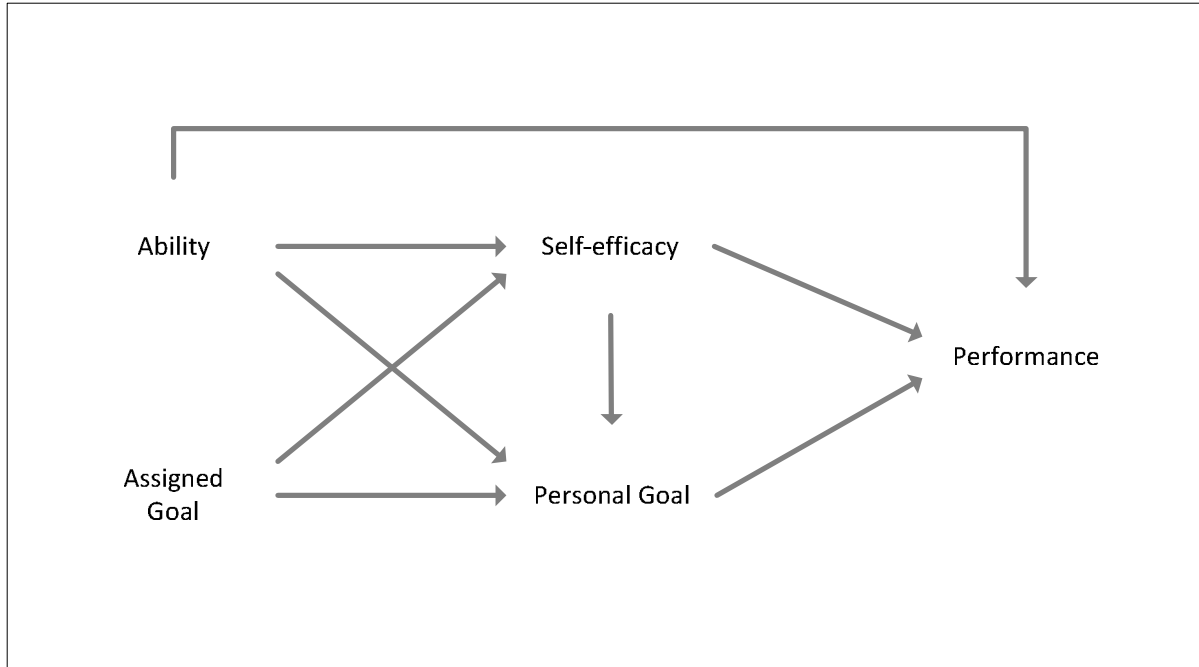
Figure 2. Goals, self-efficacy, and performance. Adapted from "Self-regulation through goal setting," by G. P. Latham and E. A. Locke, 1991, *Organizational Behavior and Human Decision Processes, 50*, p. 221. Reprinted with permission from Academic Press, Inc.
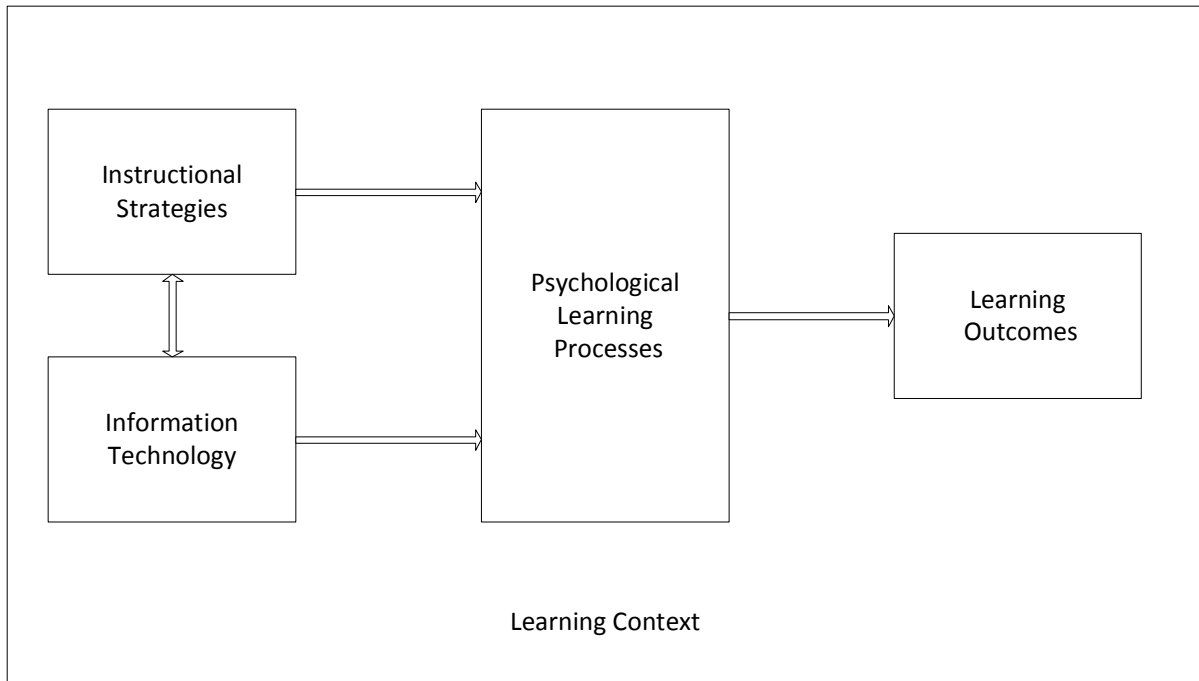
Figure 3. Framework for technology-mediated learning. Adapted from "Research commentary: Technology-mediated learning – A call for greater depth and breadth of research," by M. Alavi and D. E. Leidner, 2001, *Information Systems Research, 12*, p. 5. Reprinted with permission from INFORMS.

Technology Factors (Research Question 1)

Davis (1989) proposed a technology acceptance model whereby the perceived usefulness and perceived ease of use of information technology predict intention to use, and intention to use predicts usage behavior. Subsequently, Goodhue and Thompson (1995) added to the utilization literature by describing task-technology fit (TTF) as "the degree to which a technology assists an individual in performing his or her portfolio of tasks" (Goodhue & Thompson, 1995, p. 216). They discussed how perceptions of task-technology fit affect the utilization of technology through precursors such as the expected consequences of utilization (perceived usefulness). Dishaw and Strong (1999) then combined these two models to propose a direct relationship between TTF and actual tool usage. Eccles (2005) described usefulness or "utility value" as a component of subjective task value, which in his expectancy-value theoretical model is directly related to achievement-related choices and performance. Bandura (1997) also addressed the roles of expectancy, motivation, and outcomes in his discussion of expectancy-value theory. Citing a study on monetary incentives, he also proposed a role for perceived efficacy as a mediator of the relationship between incentives and performance. Thus, in this study, task-technology fit was hypothesized to be related to perceived usefulness, time on task for those parts of the training that are expected to be technology enabled (e.g., practice exams, assigned readings and exercises), and time spent on practice exam questions. In addition, perceived usefulness was hypothesized to be related to training self-efficacy.

Instructional Factors (Research Questions 2, 3, and 6)

Bandura (1997) defined self-efficacy as a concept formed by the "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainment" (p. 3). The four sources of information that an individual employs to develop this confidence

7

include enactive mastery experiences, vicarious experiences, verbal persuasion, and psychological or affective states. Bandura described the first source, enactive mastery experiences (i.e., doing it yourself), as the most influential, and noted that successes build a strong belief in personal efficacy, whereas failures erode it. Vicarious experience refers to modeled attainments such as those in which an individual has no absolute measure of achievement, but rather must judge his/her relative performance against the performance of others. Again, outperforming peers raises self-efficacy, while being surpassed reduces it.

Verbal persuasion is the third source of information used in forming self-efficacy beliefs (Bandura, 1997). When significant individuals express their belief in a person's ability to be successful, it helps to strengthen that person's sense of efficacy. It can also be a source for maintaining that self-efficacy when the person experiences difficulties or failures. The final sources of efficacy information are an individual's own physiological and affective states. As Bandura (1997) noted, "People often read their physiological activation in stressful or taxing situations as signs of vulnerability to dysfunction," and as a result, "the fourth major way of altering efficacy beliefs is to enhance physical status, reduce stress levels and negative emotional proclivities, and correct misinterpretations of bodily states" (p. 106). The four sources of information represent different factors working in different ways, Bandura indicated that these four factors are interrelated and are evaluated and given varying weights in the formation of efficacy beliefs for different domains. Training activities engage the first three factors, and time on task was proposed as an indicator of that training engagement. Thus time on task, a Chickering and Gamson (1987) instructional factor, was proposed to be related to PHR training self-efficacy. In addition, Bandura described enactive mastery experiences as the most influential of the four sources of self-efficacy because "they provide the most authentic evidence

of whether one can muster whatever it takes to succeed" (p. 80). Since it measures a key form of enactive mastery experience in this domain, time spent on practice exam questions, a factor suggested by prior certification exam research, was proposed to be related to PHR training self-efficacy.

Another instructional factor from Chickering and Gamson (1987) that was hypothesized to affect learning in PHR training was feedback, which was described by Bandura (1997) as an input to both enactive mastery (performance information) and verbal persuasion (efficacy appraisals). These two factors then serve as sources of self-efficacy. Also, building on goal theory and the mechanisms that have been shown to affect goal choice, a relationships is hypothesized between feedback and self-set goals. Locke, Shaw, Saari, and Latham (1981) suggested the role of feedback as a necessary component, along with goals, for performance improvement. Thus, feedback was hypothesized to affect both branches of the motivation hub model with relationships to PHR training self-efficacy and self-set goals.

Goals are also situational, and an individual might set higher goals for one activity than for another (Locke, 1991). Research supporting this theory has shown that goals are affected by a number of items, including such things as normative information, role models, competition, and pressure. The most direct way of affecting goal choice, however, is through its assignment by an authority figure (Latham & Locke, 1991). PHR examinees who prepare for the exam by participating in instructor-led training are influenced by an authority figure, the instructor, and this same interaction does not occur for those who select self-managed training formats. For this reason, it was hypothesized that instructor expectations, a third Chickering and Gamson (1987) factor, and type training were related to self-set goals.

Shiffrin and Atkinson (1969) described the processes for storing and retrieving information from long-term memory, its structure, and its relationship to short-term memory. In their model, rehearsal was one of the control mechanisms that influenced storage. Specifically, more cycles through short-term memory (rehearsal) increased the proportional amount of information stored in long-term memory, thus leading to improved performance. Other psychological researchers investigated and described the effect of spaced, or distributed practice, which as compared with mass practice, was proposed to increase learning. As Melton (1970) described it, "repetition improves remembering" (p. 606). The theoretical concept of spaced practice has been supported by a number of studies, summarized by Donovan and Radosevich (1999), who noted in their meta-analysis that "spaced practice was significantly superior to massed practice in terms of task performance" (p. 799). Based on this research, time spent on practice exam questions was hypothesized to be related to PHR certification exam score.

Rosenthal (1994) developed the theory of interpersonal expectancy effects based on work he had done in the field of experimenter outcome-bias. When applied to the fields of education and training, the theory (also referred to as the Pygmalion effect) proposed that a learner's performance could be affected by an instructor's expectations. The theory has been challenged since it was first proposed in the 1960s, but it has also been supported in numerous studies. In both a meta-analytic review by McNatt (2000) and a literature review by Murphy, Campbell, and Garavan (1999), the researchers reported studies that evaluated training situations and exam performance, finding support for the theory and the positive relationship between expectations and performance. This research supports the role of type of training, one of the instructional factors suggested by prior certification exam research to have an effect on learning outcomes. Thus it was hypothesized that type of training affects PHR certification exam score.

The Motivation Hub (Research Questions 4 and 5)

While studies of academic performance have tended to use general measures of self-efficacy or confidence, there is a body of literature indicating that domain or context-specific measures are better predictors. How far to go in establishing this specificity, however, is more difficult, and different opinions can be found. For instance, Bandura (2006) noted that "scales of perceived self-efficacy must be tailored to the particular domain of functioning" (pp. 307-308). Another researcher who addressed this topic wrote as follows:

> Domain-specific assessments, such as asking students to provide their confidence to learn mathematics or writing, are more explanatory and predictive than omnibus measures and preferable to general academic judgments, but they are inferior to task-specific judgments because the subdomains differ markedly in the skills required. (Pajares, 1996, p. 1)

For this reason, the study employed two related but distinct measures of domain specific self-efficacy: PHR training self-efficacy and PHR exam self-efficacy. In the model, it was proposed that PHR training self-efficacy affects PHR exam self-efficacy.

The theory of goal setting addresses the question of why some people perform specific tasks better than others with equal ability. According to Locke (1991), "Goals affect action by affecting the intensity, duration, and direction of action" (p. 293). Goal theory attributes the performance differences to motivation and the establishment of performance goals (Latham & Locke, 1991). Locke et al. (1984) summarized the research on goal theory by saying that "in most goal-setting studies, goals lead subjects to direct their actions in line with goal requirement, to expend effort in proportion to goal difficulty, and/or to persist in a given task until the goal is reached" (p. 241). Wood and Locke (1987) also summarized prior research by noting that "the

positive effect of goal setting on task performance is an extremely well-documented finding in the work motivation literature" (p. 1014).

The two theories of self-efficacy and goal setting were combined by Locke (1991), Latham and Locke (1991), and others into a model that predicts performance. Locke described goals and self-efficacy as "the most direct and immediate motivational determinants of performance" (p. 293). He called this construct the "motivation hub." In his model, both self-efficacy and personal goals predict performance, and also self-efficacy directly affects personal goals. Thus, in the proposed model, both PHR exam self-efficacy and self-set goals were hypothesized to be related to PHR certification exam score. Also, PHR training self-efficacy was hypothesized to be related to self-set goals.

Mediation (Research Questions 7 and 8)

The model also proposed a set of relationships that suggest mediation effects are involved. Baron and Kenny (1986) characterized a mediator as being a variable that accounts for the relation between a predictor and an outcome. The existence of mediation was suggested by the set of relationships hypothesized to exist between independent, dependent, and the proposed mediator variables. First, time on task and perceived usefulness were proposed as mediators of the relationship between task-technology fit and PHR training self-efficacy. Also, self-set goals was proposed as a mediator of the relationship between instructor expectations and PHR certification exam score.

<div align="center">Purpose of the Study</div>

The purpose of this study was to investigate the relationships between two technology variables (task-technology fit and perceived usefulness), five training variables (time on task, time spent on practice exam questions, feedback, instructor expectations, and type of training),

three motivation variables (PHR training self-efficacy, PHR exam self-efficacy, and self-set goals) and the dependent variable of PHR certification exam score (see Figure 1). It also tested whether three of these variables act as mediators.

PHR certification exam score was operationalized as an objective measure based on the respondent's self-reported PHR exam score. The PHR training self-efficacy, PHR exam self-efficacy, and self-set goals variables were measured using previously employed constructs. The 8-item training self-efficacy scale that was employed was based on research described by Pintrich, Smith, Garcia, and McKeachie (1993) and Duncan and McKeachie (2005), as modified to gauge the respondents' confidence in their PHR exam preparations. The PHR exam self-efficacy measure was derived from an academic problem-solving self-efficacy scale developed by Bandura (2006). The goals measure was based on research done by Chen, Gully, Whiteman, and Kilcullen (2000). Type of certification training was a categorical variable with seven values (live classroom-presentation course, online Internet-based course, self-study course/computer, self-study course/text, regional "crash course," flash cards, and no formal preparation). It was based on a similar approach used by Grange et al. (2003). The measure of time on task was adapted from a 9-item instrument developed by Biderman, Nguyen, and Sebren (2008). Time spent on practice exam questions was a self-reported, scaled variable. The feedback measure was adapted from a 6-item scale developed by Oberst (1995). The measure of instructor expectations was an adaptation of a 5-item instrument developed by Lee and Bobko (1992). The perceived usefulness scale was a 3-item measure based on a scale that was validated in a study by Hu, Clark, and Ma (2003). Finally, the 12-item task-technology fit scale was an adaptation of the four subscales validated in research by Staples and Seddon (2004). Those four dimensions addressed work compatibility, ease of use, ease of learning, and information quality.

Research Questions and Null Hypotheses

The research questions and null hypotheses are listed below, and the hypotheses are illustrated in Figure 1.

1. Is there a relationship between task-technology fit and the variables of time spent on practice exam questions, time on task and perceived usefulness, and also between perceived usefulness and PHR training self-efficacy?

   Ho1a: For PHR certification examinees, a self-report measure of task-technology fit will not be positively related to time spent on practice exam questions, as measured by a self-report instrument.

   Ho1b: For PHR certification examinees, a self-report measure of task-technology fit will not be positively related to time on task, as measured by a self-report instrument.

   Ho1c: For PHR certification examinees, a self-report measure of task-technology fit will not be positively related to perceived usefulness, as measured by a self-report instrument.

   Ho1d: For PHR certification examinees, a self-report measure of perceived usefulness will not be positively related to PHR training self-efficacy, as measured by a self-report instrument.

2. Is there a relationship between the factors of time spent on practice exam questions, time on task, and feedback and the variable PHR training self-efficacy?

   Ho2a: For PHR certification examinees, a self-report measure of time spent on practice exam questions will not be positively related to PHR training self-efficacy, as measured by a self-report instrument.

   Ho2b: For PHR certification examinees, a self-report measure of time on task will not be positively related to PHR training self-efficacy, as measured by a self-report instrument.

   Ho2c: For PHR certification examinees, a self-report measure of feedback will not be positively related to PHR training self-efficacy, as measured by a self-report instrument.

3. Is there a relationship between the factors of type of training, instructor expectations, and feedback and the variable self-set goals?

   Ho3a: For PHR certification examinees, a self-report measure of type of training will not be positively related to self-set goals, as measured using a self-report instrument.

   Ho3b: For PHR certification examinees, a self-report measure of instructor expectations will not be positively related to self-set goals, as measured using a self-report instrument.

Ho3c: For PHR certification examinees, a self-report measure of feedback will not be positively related to self-set goals, as measured by a self-report instrument.

4. Is there a relationship between PHR training self-efficacy and measures of self-set goals and PHR exam self-efficacy?

Ho4a: For PHR certification examinees, a self-report measure of PHR training self-efficacy will not be positively related to self-set goals, as measured by a self-report instrument.

Ho4b: For PHR certification examinees, a self-report measure of PHR training self-efficacy will not be positively related to PHR exam self-efficacy, as measured by a self-report instrument.

5. Is there a relationship between the variables of PHR exam self-efficacy and self-set goals and the measure PHR certification exam score?

Ho5a: For PHR certification examinees, a self-report measure of PHR exam self-efficacy will not be positively related to self-reported PHR certification exam score.

Ho5b: For PHR certification examinees, a self-report measure of self-set goals will not be positively related to self-reported PHR certification exam score.

6. Is there a relationship between the factors of time spent on practice exam questions and type of training and the measure PHR certification exam score?

Ho6a: For PHR certification examinees, a self-report measure of time spent on practice exam questions will not be positively related to self-reported PHR certification exam score.

Ho6b: For PHR certification examinees, a self-report measure of type of training will not be positively related to self-reported PHR certification exam score.

7. Is the impact of task-technology fit on PHR training self-efficacy mediated by time on task and perceived usefulness?

Ho7: For PHR certification examinees, the effect of task-technology fit on self-efficacy will not be mediated by time on task and perceived usefulness.

8. Is the impact of instructor expectations on PHR certification exam score mediated by self-set goals?

Ho8: For PHR certification examinees, the impact of instructor expectations on PHR certification exam score will not be mediated by self-set goals.

Limitations

1. Response rates were dependent on the researcher's ability to identify, contact, and obtain responses from PHR exam test-takers.

2. Response error and bias were difficult to control given that an online measurement instrument was used to collect data.

3. Respondents may not have been honest due to the nature of the information requested, which included certification test score ranges.

4. Random selection and assignment was not used, and therefore, external validity is affected.

5. Respondents may not have accurately recalled and reported responses to the items contained in the instruments measuring factors such as their self-efficacy and self-set goals.

6. The research design of this study did not completely control for the effects of other variables, and it did not include longitudinal data. For this reason, the ability to draw conclusions regarding causation is limited.

Delimitations

1. The study did not consider any certification programs other than the PHR exam.

2. The study did focus on the effect of training on certification testing results. It was not designed to measure the impact of the training on skills transfer.

3. A limited set of demographic data was obtained in the study, including years in the workforce, years in an HR position, industry, HR subfield, age range, gender, and level of formal education. The demographics of the study group were not evaluated in this study.

4. The study employed a self-reported measure of certification test results. The study did not validate certification test results independently with the HRCI (the PHR certification exam testing body).

5. The self-efficacy scales utilized were based on previously validated scales for learning and exam self-efficacy that were modified to address the specific domain of PHR examination self-efficacy. General self-efficacy was not evaluated in the study.

6. The training characteristics that were correlated with changes in self-efficacy and self-set goals were factors associated with the respondents' self-reported classroom, online, or self-managed training preparations. No attempt was made to evaluate vicarious, informal, or job-related learning factors.

Definition of Terms

Certification: Professional competency program that is regulated and administered by a professional body and focuses on measuring competencies and policing a profession (Wiley, 1995).

Enactive mastery: One of the four primary sources of self-efficacy as defined in Bandura's social-cognitive theory (Bandura, 1997).

Feedback: In this study, feedback refers to providing feedback during the training program. It is one of the seven instructional best practices described by Chickering and Gamson (1987).

Instructor expectations: For this study, instructor expectations refers to setting high expectations for trainees. It is one of the seven instructional best practices described by Chickering and Gamson (1987).

Perceived ease of use: Perceived ease of use is a factor from Davis's (1989) technology acceptance model. In this study, it refers to the perceived ease of use for the technology that is facilitating learning tasks.

Perceived usefulness: Perceived usefulness is a factor from Davis's (1989) technology acceptance model. It is used in this study to measure the technology's perceived usefulness in performing learning activities.

PHR Exam: Certification exam that, once passed, allows someone to use the Professional in Human Resources (PHR) designation (Forman & Cohen, 1999).

Professional in Human Resources (PHR): Certification program for HR professionals administered by Human Resource Certification Institute, or HRCI (Forman & Cohen, 1999).

Self-efficacy: Task-specific self-confidence. A key component of Albert Bandura's social-cognitive theory (Bandura, 1997).

Self-set goals: Specific levels of desired performance, in this case PHR exam performance level, that are defined by an exam test-taker (Phillips & Gully, 1997).

Task-technology fit: In this study, task-technology fit (Goodhue & Thompson, 1995) refers to the fit of the training technology to the task of preparing for the PHR exam.

Time on task: Time on task is defined as time spent performing on-task behavior, which for this study entails behavior appropriate to the task of preparing for the PHR exam (Karweit & Slavin, 1982). It is one of the seven instructional best practices described by Chickering and Gamson (1987).

Type of training: Identifies the delivery mode used for the training (Grange et al., 2003).

Verbal persuasion: One of the four primary sources of self-efficacy as defined in Bandura's social-cognitive theory (Bandura, 1997).

Vicarious experience: One of the four primary sources of self-efficacy as defined in Bandura's social-cognitive theory (Bandura, 1997).

Summary

This chapter provided background on the PHR certification exam and its use in establishing professional competency in the field of human resources. It also identified a need to better understand the factors associated with achieving a positive outcome on the PHR exam. The chapter also provided a theoretical framework and presented the purpose of the study, focusing on the topics of technology factors, training factors, self-efficacy, and self-set goals. Finally, the chapter presented research questions and associated hypotheses, limitations, delimitations, and relevant terms used in the study. Chapter 2 provides a review of the relevant literature.

CHAPTER 2

LITERATURE REVIEW

Introduction

This literature review considers how task-technology fit and the technology acceptance factor of perceived usefulness affect the utilization of technology and performance using that technology. Next, it considers how five training factors, time on task, time spent on practice exam questions, type of certification training, instructor expectations, and feedback, are related to self-efficacy, self-set goals, and outcomes. The review then explores research describing how self-efficacy has been studied in an academic setting and the role that domain specificity plays when measuring self-efficacy. It also examines research to determine how self-efficacy and self-set goals impact performance. Finally, the review considers research on the topic of mediation and the role of mediators in this study.

Technology Factors (Research Question 1)

Task-technology Fit

Goodhue and Thompson (1995) proposed the construct of task-technology fit (TTF) as an addition to utilization in an overall model they referred to as the "technology-to-performance chain." They tested a subset of that model, including the relationships between TTF and both utilization and performance, in two organizations with 26 departments that utilized 25 different technologies. They found contradictory results for the link to utilization, but support for the relationship to performance. In their study, TTF explained 14 % of the variance in performance, as measured by perceived effectiveness, productivity, and job performance. Staples and Seddon (2004) applied four dimensions of TTF (work compatibility, ease of use, ease of learning, and information quality) to an analysis of two groups. One was a group of university librarians using

library support systems, and the second was a group of students who used productivity tools in their courses. Their results demonstrated a large positive association between TTF and performance (path coefficients of .77 for the first group and .64 for the second). They also found a significant path relationship from TTF to expected consequence of use (.80 and .66). Their instrumentation of the expected consequence of use construct included adaptations of items from Davis's (1989) perceived usefulness scale.

In another learning environment, McGill and Klobas (2009) applied the technology-to-performance chain constructs to study the impact of using a learning management system (LMS). What they found was that task-technology fit for the LMS was strongly related to perceived impact on learning, but had less effect on actual student grades (path coefficients of .53 and .12, respectively). Similarly, Yu and Yu (2010) used TTF to explore how the perceptions and technology characteristics of a university e-learning environment affected graduate and undergraduate students' utilization of that technology. These researchers tested utilization as the outcome variable but did not investigate the effect of the technology on academic performance. In this case, their TTF construct was found to be moderately related to utilization of the Web site systems by the learners (path coefficient of 0.30).

Perceived Usefulness

In his technology acceptance model, Davis (1989) hypothesized that perceived usefulness and perceived ease of use are predictors of intention to use technology, and, as a result, they affect usage behavior. As a part of his investigation, he also performed two studies to validate the constructs of perceived usefulness and perceived ease of use and to test the reliability of both scales. In a larger study that investigated the factors influencing perceived usefulness and perceived ease of use, Venkatesh and Davis (2000) performed four longitudinal studies, each

testing these variables over three periods of time. In all four cases, perceived usefulness was strongly related to intention to use (β = .44 to .64 over the 12 results). As predicted by the theory, intention to use the technology then predicted usage behavior. When all four studies were pooled over the three measurement time horizons of the longitudinal study, SEM results showed a path coefficient of .55 from perceived usefulness to intention to use. Then the path coefficient of the pooled model and data between intention to use and usage behavior was .52.

Yi and Hwang (2003) applied this model in a study of the Blackboard system in an academic setting. In this case, the path coefficient from usefulness to behavioral intention was .46. However, the path coefficient from behavioral intention to use was only .19. In another study in an academic setting, teachers' use of PowerPoint was evaluated both prior to and after training (Hu et al., 2003). What these researchers found was that the path coefficient from perceived ease of use to perceived usefulness was .57 after the training. As compared with the above studies, this is a new result since they did not investigate the ease of use to usefulness link. Also, the path coefficient for the link between perceived usefulness and intention to use was a strong .83 after training. Cheung, Li, and Yee (2003) tested the relationship between perceived usefulness of a multi-media learning system and learning self-efficacy in the teaching of database skills. They found that for MIS students (but not computer science students), perceived usefulness, as well as perceived ease of use, were significantly related to improved learning self-efficacy.

<div align="center">Instructional Factors (Research Questions 2, 3, and 6)</div>

Time on Task

Time on task is a variable that has been widely studied in both educational and training settings. Karweit and Slavin (1982) defined on-task behavior as "behavior appropriate to the

<div align="center">22</div>

task at hand" (p. 846) and further noted that the definition of appropriate behavior depends on both the task and the specific rules defining the learning environment. In their study, the learning environment was the classroom, and time on task was measured through observation. What these researchers found was that time on task did predict learning as measured by posttest results. However, they also noted that factors such as adjusting the definition of on-task and off-task behavior and modifying the length of observation both affected the magnitude of the correlations.

In a study of online training in an organizational setting, Brown (2001) tested the effects of both time on task and practice level on learning. In this case, time on task was recorded by the training software and included the total time spent on learning modules. Practice level was also captured by the software and was based on the specific amount of time trainees spent on practice activities during their training. The researcher did not include quizzes in the total since they were required of all participants and thus were not a variable factor. What Brown found was that both practice level ($\beta = .33$) and time on task ($\beta = .18$) predicted knowledge gain as measured by a pretest-posttest difference score. Also significant was the result that the individual differences he tested were weak predictors of both practice level and time on task.

Finally, both Oberst (1995) and Robertson, Grant, and Jackson (2005) investigated the seven instructional factors suggested by Chickering and Gamson (1987) in classroom settings. In Oberst's study, a canonical correlation with four of those seven instructional factors, along with four other confounding factors, was found to significantly differentiate between high achievers and at-risk students, as measured by mean differences in GPAs. Specifically, time on task had the largest standardized canonical discriminant function coefficient (.64). Thus, time on task was found to be the largest predictor of high achievement in the academic environment

studied by the researcher. Another interesting outcome was the large correlation he found between time on task and feedback (.51). Perhaps as a result, the stepwise analysis he performed failed to find that feedback contributed significantly to the predicative ability of the group of factors. In the study by Robertson et al., the researchers compared graduate courses delivered online and those taught in the classroom and asked students how each of the factors contributed to their learning success. Time on task was found to be the only one of the seven factors with a significant mean difference.

Time Spent on Practice Exam Questions

A number of studies have examined the correlation between factors associated with certification exam training and subsequent certification test results. One factor mentioned in some of these prior certification exam studies is the impact of practice exams. For instance, in a study of the Texas teacher certification exam performed by Simonsson et al. (2000), ExCET practice exam scores were found to be positively correlated with teacher certification exam results. Correlation results showed that practice exam results accounted for 25% of the variance in subsequent ExCET results. Also, a study by Harrington et al. (1992) showed that practice exam results were positively correlated with teacher math skill certification test results. In this study, the sample or practice test accounted for 39% of the variance in the criterion variable, the actual test score. Finally, Dunn and Hall (1984) demonstrated that for those seeking the Certified Public Accountant (CPA) certification, hours of self-study was a variable positively related to certification outcomes. This factor had the third highest level of correlation, after GPA and SAT® scores (College Board, New York, NY, www.collegeboard.org), accounting for between 4% and 7% of the variation in the scores for each of the four sections of the CPA exam.

Type of Certification Training

There are several ways of preparing for certification testing. The options include formal test preparation classes and different types of self-study programs. In several professional exam settings, studies have found that test preparation programs are an effective way of improving test performance. In the realm of the CPA exam, several studies have shown that taking a review course was positively correlated to test outcomes. In fact, in one study by Grant et al. (2002) the authors quantified the benefits derived from review courses, suggesting that they were the most efficient mechanisms. They suggested that a CPA exam taker derived the same improvements in outcome from two thirds of a review class as they did from 22 hours of additional course work. In addition, Titard and Russell (1989) tested the pass rates of those CPA examinees who participated in formal supplementary study and those who did not. The Z scores calculated between the two groups for each section of the exam were so high that the authors concluded with over 99% confidence that supplementary study was a positive factor associated with passing the exam.

Finally, in the other business-related certification example cited earlier, the Certified Financial Planner™ exam (CFP Board, Washington, DC, www.cfp.net), Grange et al. (2003) looked for the same effect. In addition to noting that almost all examinees took a preparation course, they found that preparing for the Certified Financial Planner exam by attending a live presentation or lecture or by reviewing a textbook increased an individual's pass rate by 1.7 times as compared to preparation using computer-based materials.

Benefits from preparation programs have not been found in all settings, however. Ryan, Ployhart, Greguras, and Schmit (1998) examined the impact of participation in a firefighter exam preparation class, as compared with other factors such as ability, demographics, "dispositions"

25

(e.g., ambition), and study skills. They hypothesized that the primary benefit of such a program should be improved performance on the test, but the authors also expected additional potential benefits, including increased motivation and decreased test anxiety, that would in turn affect test performance. Their research did not find that the test preparation program significantly improved test performance and motivation, nor did it reduce test anxiety. However, they noted that this outcome could perhaps be due to the short duration of the program (only a few hours in length), and the lower initial ability level in the group of participants who chose to attend.

Instructor Expectations

The role of expectations in performance has been studied over a number of years. Locke et al. (1984) performed a controlled experiment with college students involving task-based training. In their experiment, which extended over seven trials, they taught one group an involved task-solving procedure, another had an opposite form of training, and the third was a control. Then they tested the students' performance in conditions both with and without assigned goals. The researchers found that both the assigned goals and the students' self-set goals were related to performance, and this extended to subsequent trials. In addition, those with assigned goals set higher performance goals for themselves, and their self-set goals remained higher than the other group over the subsequent trials, although the difference declined over time. Similarly, those with assigned goals performed better than those without assigned goals, although this difference also declined in subsequent trials.

Latham and Locke (1991) summarized their previous research by suggesting that "the correlation between assigned and (subsequently) self-set goals is around .50, indicating that goal assignment does affect choice although it obviously does not totally determine choice" (p. 220). Klomegah (2007) investigated the linkage between assigned and self-set goals and their

relationship with academic performance. He found that the correlation between the two was significant (.42) and that self-set goals was a significant predictor of course grade, explaining 23% of the performance variation, and that assigned goals explained 17% of the variation in course grade. Finally, qualitative research by Dennen (2005) that evaluated various instances of online instruction demonstrated a different aspect of the expectation linkage to performance. The researcher found that the lack of explicit expectations regarding the class itself negatively affected student participation in the online environment. Lack of student participation could obviously have a downstream performance effect, and it suggests another dimension of instructor expectations.

Feedback

Latham and Locke (1991) discussed the role of feedback in goal setting. They noted that goal setting mediates the relationship between feedback and performance. Also, feedback acts as a moderator in the relationship between goals and performance. They concluded that taking these two outcomes together, "goals and feedback are more effective in motivating high performance or performance improvement than either one separately" (pp. 225-226). The researchers summarized the results of 33 studies by noting that in 17 cases goals and feedback together led to better performance than goals alone; in 20 cases the combination was better than feedback alone; and in only 3 instances was it worse than either feedback or goals separately.

In an earlier summarization of research on goal setting conducted over the previous decade, Locke et al. (1981) reported 10 studies that investigated goal setting and feedback. First, they considered 4 studies that compared task performance for three combinations of goals and feedback: specific, hard goals with feedback, no specific goals with feedback, and no specific goals with no feedback. They concluded that feedback does not improve performance

27

independently of goals. In the next 4 studies they reviewed combinations of specific, hard goals with and without feedback, and no specific goals with no feedback. Here, they concluded that goals with feedback also do not improve performance. In the last 2 studies they described, the researchers compared task performance in a specific goal case, both before and after receiving feedback. In these cases they concluded that "providing explicit or frequent feedback clearly facilitated performance" (p. 135). Based on this combination of results across multiple studies, the conclusion that they proposed was the joint action mechanism discussed above. Thus, specific, hard goals and feedback must be combined to improve task performance. Consistent with these results, Bandura and Cervone (1983) found that the combination of feedback and goals improved task performance above either individually. Schunk (1990) also reported a study in which the combination improved performance in an academic setting.

In their review of prior empirical studies discussed above, Latham and Locke (1991) also noted that positive feedback tended to raise self-efficacy. Schunk (1990) discussed a number of his own studies in classroom settings that indicated both self-efficacy and performance are improved when students receive feedback. Bandura and Cervone (1983) provided different combinations of feedback to students performing a new physical task they had to learn and found that task efficacy was related to the feedback they received rather than their actual performance. This is consistent with a discussion by Bandura (1997), who commented on Schunk's studies, noting that "ability feedback in the early stages of skill development has an especially notable impact on the development of a sense of personal efficacy" (p. 102).

The Motivation Hub (Research Questions 4 and 5)

Self-efficacy and Taking Tests

As Fletcher (2005) noted, self-efficacy scales have been developed for a number of specific task domains, including academic, computing, and Internet use. In the academic domain, Tannenbaum, Mathieu, Salas and Cannon-Bowers (1991) found that self-efficacy was related to how U.S. Naval recruits performed on three academic tests over the course of their 8-week training program. Their analysis demonstrated that test performance was positively related to academic self-efficacy ($\beta = .272$). Also, Finney and Schraw (2003), in their study of self-efficacy in regard to statistics knowledge and test-taking, found a positive relationship between students' current statistics self-efficacy at the end of a statistics course and their total course percentage (performance). The researchers' analysis found the magnitude of this correlation to be significant, with $r = .496$. In their study of student self-efficacy in the domain of mathematics performance, Pajares and Kranzler (1995) reported that " students' self-efficacy beliefs about their math capability had strong direct effects on math anxiety and on mathematics problem-solving performance even when general mental ability was controlled" (p. 437). In their study, the correlation between math self-efficacy and math performance had an $r = .64$.

Domain-specificity in Measuring Self-efficacy

The specific way in which self-efficacy is measured can have an impact on the effect size of the linkage between self-efficacy and performance. As an example, Lane, Lane, and Kyprianou (2004) evaluated three types of self-efficacy measurement in a classroom setting. Two of the self-efficacy measures related to competence (maintaining motivation and coping with demands of the program), and one related to performance (passing at the end of the program). Only the third measure was a statistically significant predictor of performance ($p <$

29

.001), and it accounted for 10.4% of performance variance. Lane et al. also summarized the literature, noting that the following factors tended to maximize the effect size: (a) "knowledge of the task to be performed", (b) "short time lag between self-efficacy ratings and task performance", (c) "self-efficacy measures and performance that lie in the same behavioral domain", and (d) "specific tasks" (p. 248). Conversely, weaker effect size was associated with general tasks, task complexity, and especially, "complex tasks involving heavy demands on knowledge, cognitive ability, and persistence" (p.248).

Nielsen and Moore (2003) demonstrated the value of domain-specificity in self-efficacy measurement in their analysis of the Mathematics Self-Efficacy Scale (MSES). They tested the correlation of the MSES with mathematics tests scores and evaluated how the level of context specificity affected the measurement of self-efficacy. They did this by administering the MSES instrument in two contexts, the classroom and a test setting. Although the items were the same, the context was altered for the two self-efficacy scales. The class and test versions of the MSES scale were found to have positive correlations with last grade (.51 and .58) and desired future grade (.52 and .58). When examining these results, they discovered that for mathematics self-efficacy, this distinction was important in at least two ways. Overall, the students they administered the exam to were found to have statistically significantly higher mathematics self-efficacy in a classroom context than when the context was a test setting. In addition, the difference between the two measures of self-efficacy was even more pronounced when the students were stratified based on their total level of self-efficacy across both domains and the lower efficacy group was analyzed.

Another relevant example of domain or context specificity can be found in the distinction between process and outcome self-efficacy. Mone (1994) pointed out that "although they are

likely related, process and outcome self-efficacy may each have predictive abilities for subsequent cognitions (e.g., goals) and performance" (p. 517). His study evaluated the relationship between student outcome (grade) self-efficacy, process (academic) self-efficacy, and performance. The results supported the hypothesis that outcome self-efficacy (r = .31, .32, and .34 for three test iterations) predicted performance better than process self-efficacy (r = .20, .12, and .19 for the same three test iterations). He likened this to the difference between general and task-based self-efficacy and their predictive abilities, noting that the predictor should be closely related to its object. The conclusion to be drawn from these studies is that context must be considered when choosing a self-efficacy scale or instrument in order to ensure measurement validity.

Self-efficacy, Self-set Goals and Performance

In an exploratory study investigating the effects of goals, task strategies and self-efficacy on task performance, Locke et al. (1984) showed that both goals and self-efficacy positively affected task performance in a group of undergraduate students. The path analysis results they reported showed a path coefficient of .20 for the self-efficacy to performance relationship and .24 for the goal to performance relationship. Performance was based on a cognitive task (describing uses for common objects), and was not specifically based on testing. However, the researchers did evaluate how strategy training, ability, and assigned goals affected the variables of self-efficacy, self-set goals, and performance. They performed multiple trials to see what effect previous goals and task performance would have on the variables in subsequent iterations. One finding was that a relationship existed between student self-efficacy and their self-set goal choice.

31

These results were subsequently repeated and incorporated by Locke (1991), Latham and Locke (1991), and others into a model that combined self-efficacy and self-set goals as predictors of performance. Locke pointed out that goals and self-efficacy "are considered to be the most direct and immediate motivational determinants of performance" (p. 293), with cognitive factors such as ability, task knowledge, and strategy also having relevancy. He went on to "call the goal/self-efficacy/performance linkages the *motivation hub*" (p. 296).

A number of studies of the motivation hub have been conducted in which the objective was to measure its effect on academic performance. For instance, in a study of undergraduate students, Wood and Locke (1987) found the relationship between self-efficacy and academic performance to be significant, even after controlling for ability. Regarding student self-efficacy, they found a distinction between the measurements of self-efficacy magnitude (could they achieve a specific outcome) and its strength (mean confidence level). In particular, strength was more highly correlated with performance than was magnitude. The weighted mean correlation for the three studies the researchers reported was .27 for the relationship between self-efficacy strength and performance, as compared with .18 for self-efficacy magnitude and performance. In structuring the self-efficacy measures, the authors followed the task-specificity suggestions of Bandura and others and broke their self-efficacy measure into seven specific course subject/task areas. They also discussed their findings of the relationship between a four-component goal construct and performance by noting that grade goals had a significant effect on performance (weighted mean correlation of .42 for the three studies).

Phillips and Gully (1997) studied undergraduate students and how self-efficacy and grade goals affected exam performance. In addition to the motivation hub variables, they also investigated ability, locus of control, learning goal orientation, and performance goal orientation.

Their results showed that self-efficacy had a direct positive effect on academic performance (path coefficient of .23), and that "self-efficacy contributes to performance prediction above and beyond both ability and goal prediction" (p. 797). Similarly, self-set goals had a positive effect on performance, with a path coefficient of .21. Regarding the other variables, Phillips and Gully found that learning goal orientation and locus of control both positively affected self-efficacy and that performance goal orientation negatively affected it. In addition, the relationships between ability and both self-efficacy and performance were both found to be positive, but the effect on self-set goals operated through self-efficacy.

In a final example, Chen et al. (2000) studied students in an undergraduate classroom setting where performance was defined in terms of exam grade. Like the other researchers, they investigated the motivation hub variables of task-specific self-efficacy and goals and again found that these variables predicted performance. In this case, Chen et al. also studied the variables of cognitive ability, learning goal orientation, performance goal orientation, state anxiety, and general self-efficacy. The authors performed two similar studies that used different measures of goal orientation and state anxiety. The reported path loadings were .21 and .25 for the goals to performance relationship and the specific self-efficacy to performance path loadings were .13 and .25. Some results they found regarding the additional variables included a relationship in which general self-efficacy predicted task-specific self-efficacy, unlike the suggestions of Bandura (1997). In addition, task-specific self-efficacy did not fully mediate the relationship between cognitive ability and goals, unlike the results of the Phillips and Gully (1997) study. Chen et al. also found that learning goal orientation affected task-specific self-efficacy through general self-efficacy and that performance goal affected it through state anxiety, supporting the distinction between the two types of goal orientation. Finally, cognitive ability was found to be

"more strongly related to initial performance than to subsequent performance" (p. 844). The authors cited several studies supporting the suggestion that the impact of ability on performance declines over time as task mastery increases.

As has been shown, a number of studies have evaluated how self-efficacy and self-set goals are related to performance, including some where performance is defined in terms of test results. These studies have consistently supported the existence of a motivation hub construct. At the same time, the studies have presented varying results regarding the antecedents of the motivation hub factors. They have also demonstrated examples where the motivation hub variables have acted as mediators. All of these results are consistent with the purpose of the study as described earlier in Chapter 1.

<center>Mediation (Research Questions 7 and 8)</center>

In an early and widely quoted article on mediation and moderation, Baron and Kenny (1986) described a mediator as a variable that helps explain the relationship between a predictor and an outcome. They observed that "mediators explain how external physical events take on internal psychological significance" (p. 1176). MacKinnon (1994) described seven reasons for analyzing mediation effects. Two of them, checking whether a program manipulates an intervening mediator variable and using that information to improve the program, are especially germane to this study.

Frazier, Tix, and Barron (2004) described several factors that should be considered when designing mediation studies. First, "there should be an effect to mediate" (p. 126). By this, the authors mean that a relationship between the predictor and outcome variables has usually been established in previous research. In the case of this study, research regarding the relationships between the perceived ease of use component of TTF and self-efficacy (Cheung et al., 2003) and

<center>34</center>

between instructor expectations and performance (Locke et al., 1984) was discussed earlier. Frazier et al. also suggested that there should be theoretical background to support the assertion that the predictor and mediator variables are related. The relationship between task-technology fit and perceived usefulness (expected consequences of utilization) was proposed by Goodhue and Thompson (1995). Subsequently, Dishaw and Strong (1999) combined the task-technology fit and the technology acceptance models to propose a direct relationship between TTF and actual tool use. Finally, Latham and Locke (1991) described the importance of authority figures in goal setting. This concept of authority figures affecting goal setting supports the idea of a relationship between instructor expectations and self-set goals.

A final point to consider about mediation is the topic of causation. Frazier et al. (2004) pointed out that mediation implies causation, but that causal criteria must still be established. This involves demonstrating that the cause precedes the effect. Both MacKinnon (1994) and Frazier et al. (2004) described the value of longitudinal analysis of mediated relationships in helping to establish the criteria for causality.

Summary

This chapter reviewed the research that supports the use of self-efficacy and self-set goals to investigate how training affects certification outcomes. First, the factors of task-technology fit and perceived usefulness were introduced, and their effect on technology utilization and performance was discussed. Then there was an examination of how the five training factors of time on task, time spent on practice exam questions, type of certification training, instructor expectations, and feedback are correlated with outcomes, self-efficacy, and self-set goals. The chapter also considered the topics of self-efficacy in academics and the importance of domain-specific self-efficacy measures, and it reviewed literature describing how self-efficacy and self-

set goals were related to performance. Finally, the chapter considered the topic of mediation and

the proposed mediators. Chapter 3 presents the methodology used in this study.

CHAPTER 3

METHODS

Research Design

The research design for this study was correlational and involved the testing of a

mediation model. The hypothesized model was analyzed using partial least squares (PLS), a

type of structural equation modeling. As discussed in Chapter 2, while mediation can imply

causation, demonstrating causation requires evidence that cause precedes effect (Frazier et al.,

2004). Because of the complexities of sampling Professional in Human Resources or PHR®

certification (Human Resource Certification Institute, Alexandria, VA, www.hrci.org) examinees

before training, after training, and a third time after taking the exam, such an approach was not

selected for this study. The PHR certification examinee population was sampled after they had

completed both certification training and the PHR exam itself. At that time, they were asked to

respond to survey questions about how they felt after the training. Thus, no pretraining data

were captured. As a result, the ability to draw conclusions regarding causation was limited.

Population

The target population for this study included recent PHR certification examinees who had

taken a formal PHR examination preparation class as well as those who used other exam

preparation approaches. One reason for targeting this combined population was to ensure that a

range of certification training approaches was represented in the sample. There are a number of

ways to prepare for the PHR exam, and this diversity allowed for the testing of some of the

variables in the hypothetical model, including type of certification training as well as task-

technology fit. Another reason for including those who participated in formal PHR examination

preparation classes in the target population is that two variables, instructor expectations and

feedback, were both tested using instruments that assumed an instructor-student interaction. Finally, the target population was characterized by academic, experience level, and demographic diversity. This mirrors the broader population that sits for the PHR examination.

Sample

The sampling approach for this study was convenience sampling. Examinees were contacted both by emailing colleges and universities that offered PHR examination training programs and by emailing SHRM® (Society for Human Resource Management, Alexandria, VA, www.shrm.org) chapters directly. In both cases, the individuals who were contacted were asked to forward the survey instrument to examinees who had just taken the exam in the May-June 2011 sitting. The determination of sample size in this study was affected by several factors. One factor was the nature of the hypothesized mediation model. MacKinnon (1994) noted that the statistical power of a mediation effect is less than a test of regression coefficients. As a result, the sample size required to detect mediation may need to be larger than what would be computed using traditional methods for estimating samples in a regression scenario. Also, the study was investigating a model with latent mediator variables which implied the use of structural equation modeling (SEM) for mediation analysis and not multiple regression (Iacobucci, 2008). In her discussion of mediation analysis using multiple regression and covariance based SEM, Iacobucci suggested that "SEM is more powerful than regression for all sample sizes, with the greatest differences between the techniques favoring SEM especially for small samples ($n = 30$), when the researcher can benefit from the additional compensatory power of the test" (p. 23).

As discussed in the analysis section, there are two primary approaches to performing structural equation path modeling, the covariance-based approach and the component-based or

partial least squares (PLS) approach (Hsu, Chen, & Hsieh, 2006). The PLS approach was

selected for this study for several reasons. First, the study involved the use of both formative and

reflective constructs, as discussed below under Instrumentation.   The partial least squares

method is better able to handle both types of variables (Chin, 1998; Hsu et al., 2006). In

addition, PLS does not assume variable normality (Sosik, Kahai, & Piovoso, 2009). Finally, PLS

can be used with smaller sample sizes due to its approach to analyzing path relationships (Chin,

1998; Hsu et al., 2006; Sosik et al., 2009).

When determining sample size for the study, the rule of thumb that was employed was

the widely-referenced recommendation by Chin (1998). Based on his studies of PLS, this

researcher and developer of the PLS Graph tool recommended a sample size of 10 times the

largest number of "predictors" for any individual variable in the model. In this context, the

largest number of predictors will be found to be either (a) the largest measurement equation (i.e.,

the largest number of formative indicators for any latent variable) or (b) the largest structural

equation (i.e., the dependent variable with the largest number of independent variable affecting

it) (p. 311). In the case of the model employed in this study, this calculation results in a

minimum sample size of 40 for the study.

As an additional step, sample size was computed following the recommendation of Chin

(2010) regarding the use of Cohen's (1988) power tables for multiple linear regression. Instead

of the power tables, G*Power 3.1 (Faul, Erdfelder , Buchner, & Lang, 2009) was used to show

that a sample size of 40 did indeed result in power of .80 when testing effects of the size $f^2 = .34$

(equivalent to $R^2 = .25$) with four predictor variables (IV's) in multiple linear regression. Based

on Cohen's definitions of f^2 effect sizes in multiple linear regression, power with a medium-sized

effect ($f^2 = .15$ which is equivalent to $R^2 = .13$) was also evaluated. For .05 significance and four

predictor variables, a sample size of 84 achieved approximately a .80 power with a medium effect size. Although many of the factors included in this study had exhibited large effect sizes in prior research, the larger minimum sample size of 84 was nonetheless established to optimize the power of the tests to reject the null hypotheses and to detect mediation.

Instrumentation

The selected method for performing this study was survey research. The survey was used to gather five categories of information, including demographic information, technology variables, instructional variables, motivation variables, and the outcome variable.

Demographic Information

A number of items were captured in the survey to allow for further investigation of the population sample and to aid in considerations of external validity. These demographic information survey questions appeared toward the end of the instrument but before the final section that included a set of optional instructor-led training survey items. The requested demographic data included (a) years in the workforce, (b) years in an HR position, (c) industry, (d) HR subfield describing current position, (e) age range, (f) gender, and (g) level of formal education.

Technology Variables

The technology variables in the hypothesized model were evaluated using self-reported survey variables. Task-technology fit (TTF) was measured using a 12-item instrument developed by Staples and Seddon (2004). The TTF instrument these researchers tested was a formative construct consisting of four dimensions that address work compatibility, ease of use, ease of learning, and information quality. In their study, Staples and Seddon reported coefficient

alphas of .762 for work compatibility, .840 for ease of use, .921 for ease of learning, and .822 for information quality. They also used average variance extracted and inter-correlations between the latent variables to confirm discriminant validity for these factors. The measure selected for perceived usefulness was a scale employed by Hu et al. (2003). The perceived usefulness scale was a 3-item measure which the researchers found to have a reliability of .77. The items in both the perceived usefulness and task-technology fit subscales were modified to address the domain of training technology.

Instructional Variables

The instructional variables in the hypothesized model were also evaluated using self-reported survey variables. Type of certification training was evaluated using an approach modeled on one employed by Grange et al. (2003). These researchers were able to identify a significant difference between those respondents who prepared for the Certified Financial Planner™ exam (CFP Board, Washington, DC, www.cfp.net) by attending live classroom presentations or using text-based self-study materials and those who did not. Grange et al. used an instrument with 4 values, but, based on feedback from a panel of experts, additional values were included to reflect the range of available PHR training options. Thus, the survey employed a single variable with 7 values: (a) live classroom-presentation course, (b) online (Internet-based) course, (c) self-study course/computer, (d) self-study course/text, (e) regional "crash course," (f) flash cards, and (g) no formal preparation. The time spent on practice exam questions item was a variable with 10 values ranging from 0-2 hours per week up to 18-20 hours per week, and an 11th value for greater than 20 hours per week. The number of intervals selected was intended to allow for discrimination between levels of preparation using this training activity.

Time on task was measured using a scale developed by Biderman et al. (2008). In their study, the researchers operationalized time on task as "reported study time" (p. 889). The nine-item scale they developed included eight Likert-scaled items and one continuous variable (number of hours spend studying) that was transformed to a logarithmic value during analysis. Factor analysis of the scale by the authors demonstrated that all nine items loaded on a single factor. They reported a coefficient alpha value for the scale of .79. The variable of feedback was adapted from the feedback measure developed by Oberst (1995). In his study, Oberst validated this six-item Likert-scaled measure along with the rest of the instrument he used to investigate the Chickering and Gamson best practices. From a reliability perspective, the coefficient alpha was .807 for the feedback subscale.

Instructor expectations were measured using a six-item Likert-scaled measure adapted from a study by Lee and Bobko (1992). They evaluated two subjective goal difficulty measures, one self-referenced and one that was externally-referenced. What they found was that the externally-referenced measure had the highest levels of correlation with both the actual assigned goal and subsequent performance. The researchers also reported that this measure exhibited a high degree of reliability, with a coefficient alpha of .91. Thus the externally-referenced measure was selected for use in this study, with adaptations in wording to fit the context.

Motivation Variables

The motivational variables in the study were investigated using previously validated approaches. For self-set goals, the approach was adapted from work done by Chen et al. (2000). In their research, 316 graduate students were asked to complete two items regarding their course goals. One asked for a letter grade goal (A, A-, B+, etc), and the other asked for a percentage of items correct goal (0-100%). The researchers found the two items to be highly correlated on two

sampling occasions (.78 and .70). For analysis purposes, they combined the two items into a single goal measure. Over the course of the two samples, their analysis showed the path coefficients between goals (prior to initial performance) and performance (final exam scores) to be .21 and .25. In this study, PHR exam scores were substituted for letter grades.

The instrument used to measure PHR training self-efficacy was adapted from the eight-item Self-Efficacy for Learning and Performance scale developed at the University of Michigan (Duncan & McKeachie, 2005; Pintrich et al., 1993). The items were constructed using 7-point Likert scaling. In a study of the reliability and predictive validity of that instrument, Pintrich et al. sampled a group of 380 college students, and measured their academic performance across multiple subjects. The researchers found the coefficient alpha for the self-efficacy sub-scale to be .93, indicating high internal consistency for the eight items. From a predictive ability standpoint, they found the r with the final course grade to be .41, indicating a high level of correlation. In this study, the wording of the self-efficacy scale questions was modified to reflect both the retrospective nature of the survey and the fact that the performance outcome is the PHR certification exam score instead of an academic grade.

Based on the research concerning the importance of domain specificity, the second measure of self-efficacy employed in this study was PHR exam self-efficacy. In this case, a traditional self-efficacy instrument design was employed. This scale gave different levels of performance and asked the respondent to indicate how confident they were that they could perform at each level. The responses were then summed to form a composite score. In this case, the instrument construct, format, and wording were based on an (academic) problem-solving self-efficacy scale developed by Bandura (2006).

Outcome Variable

The PHR certification outcome variable was a self-reported measure. The first section of the survey began with filter questions, confirming that the respondent had in fact completed the PHR examination. In addition, the training month and year, as well the PHR examination year and month, were also captured. These allowed for confirmation of both the sequence and proximity of the training and examination activities. The PHR exam score was then measured using a scale with ten 25-point intervals from 451 to 700 and a response if the score was below the range.

Pilot Study

Prior to data collection, a pilot study was performed to reconfirm the reliability of the individual measures and to test usability and comprehension of the instrument after the measures were combined to create a Web-based survey. That pilot study began with a review of the instrument by a panel of four experts selected based on their specific experience with the PHR or SPHR® certification (Human Resource Certification Institute, Alexandria, VA, www.hrci.org) exam. That review helped confirm content validity for the measures and also led to the addition of some new demographic items, as well as additional responses for some items such as the type of training variable. Next, the instrument was administered to a sample of 21 individuals with prior PHR exam experience to test the usability and comprehension of the instrument and to confirm the reliability of the individual measures. Pilot test members were solicited via email both from a university training program as well as from a local SHRM chapter. Those results were evaluated, and item reliability measures were found to be satisfactory and in line with the prior reported studies from which the instruments were drawn. Several questions in the survey were modified based on feedback from the pilot study participants and analysis of pilot results.

The changes were made to improve item readability, add response options, and increase response rate. Also, the sequencing and grouping of items on the instrument were adjusted based on analysis of the responses.

<center>Data Collection</center>

Data collection procedures involved contacting PHR examinees after completion of both certification exam training and the PHR examination itself. The respondents were identified and contacted via email, with the assistance of PHR examination training class administrators and SHRM chapter officers. These class administrators and chapter officers were asked to forward the email to examinees who they identified as being qualified members of the sample population. In the contact email, a link to a ZipSurvey™ (RELIANT LLC, Tulsa, OK, www.zipsurvey.com) Web site was distributed, along with a brief explanation of the study, its purpose, and the survey process.

The first page of the survey contained an IRB-required informed consent disclosure with information about the university, the researcher, confidentiality of the data, and the freedom of recipients not to respond. The survey itself took approximately 15 minutes to complete and consisted of three windows that captured demographic, technology, self-efficacy, goal, and PHR examination result data. There was no information contained in the survey that would allow for identification of respondents, and the most sensitive information in the survey was the dependent variable, PHR examination result. For those who indicated that their training was instructor-led, a fourth window was presented that asked instruction-related questions.

Survey responses were stored on a database, which was maintained on a password-protected Web site that was professionally hosted by ZipSurvey. To ensure confidentially for the respondents, the researcher had password-protected access to these data, and no individual

<center>45</center>

information was given to any of the class instructors. However, to encourage participation, a summary of the study results was offered both to those who forwarded the survey and to respondents. In addition, both groups were offered the option of enrolling in a drawing for an Apple iPad. Instructors and chapter officers who forwarded the survey enrolled in the drawing by responding to the solicitation email and confirming their interest. Respondents were provided with a link to a separate survey form where they could enter their email information to enter the drawing. They could also indicate whether or not they wanted to receive a copy of the summary study results information. Through the use of a separate response file and a separate sign-on process to access it, this information was not linked to the survey responses themselves.

Data Analysis

After completion of the survey, responses were examined to verify that they all met the criteria for age of the data, completion of both stages (completion of examination training and the exam itself), and any other anomalies in the respondents' replies (inconsistencies in exam history responses, etc.). Completeness of the data was evaluated as well to determine which respondents to include in the analysis and which responses or items to exclude due to missing or inconsistent data. Responses were also excluded if the respondent's PHR exam preparation training had occurred earlier than 1-2 months prior to taking the May exam or earlier than 2-3 months prior to taking the June exam. Finally, two tests were performed to check for common method bias in the responses (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). These included Harman's single-factor test and application of a partial correlation procedure that used a five-item social desirability scale developed by Hays, Hayashi, and Stewart (1989). This second test was included to determine if there was method bias attributable to respondents replying in a way that they felt the researcher wanted.

After completion of the data screening, SPSS® (International Business Systems, Armonk, NY, www.ibm.com) was used to initially examine scale reliabilities. Those results were compared to the original results reported by the researchers who developed the scales for task-technology fit, perceived usefulness, time on task, time spent on practice exam questions, feedback, instructor expectations, the two self-efficacy scales, and self-set goals. After completing examination of the reliability metrics, a second dataset was created which consisted of responses from individuals who had participated in instructor-led training. The second dataset was used in the third analysis step described below.

Although in MacKinnon's (2008) view the multiple regression approach detailed by Baron and Kenny (1986) may be the most widely cited and commonly used method for performing mediation analysis, it is a complicated process, with four steps and three regression equations. In the present study, there are two mediation relationships in the proposed model. Structural equation modeling can be used to accommodate such situations, and according to Iacobucci (2008), in situations involving models with latent mediator variables and latent independent variables, it is the only permissible analysis approach. In fact, she also made the point and demonstrated that, even in situations when regression could be used, "the SEM technique is the superior method on both theoretical and empirical statistical grounds" and is always appropriate (p. 20).

There are two primary approaches to performing structural equation path modeling, the covariance based approach and the component based or partial least squares (PLS) approach (Hsu et al., 2006). The PLS approach was selected for this study for a number of reasons. First, the task-technology fit instrument employed in the study was based on four subscales in a formative construct. Such constructs are better addressed by PLS, which can handle both

formative and reflective variables (Chin, 1998; Hsu et al., 2006). In addition, PLS is a

nonparametric technique that does not assume variable normality (Sosik et al., 2009). It was felt

that this could be an issue for at least one of the variables in the model (type of training). Also,

PLS is an analysis technique that can be employed with smaller samples due to the approach it

uses in analyzing path relationships (Chin, 1998; Hsu et al., 2006; Sosik et al., 2009). In the case

of this study, the complexities of indirect sampling without access to a list of examinees was

deemed likely to affect the initial number of respondents.

   As described above, the study also employed a scrubbing test that removed a number of

responses based on the time that had passed since the respondents had participated in training. In

addition, screening for incomplete data eliminated a number of responses from the analysis

process. Finally, the study anticipated two steps in the path analysis, with the second step

including responses received only from those who participated in instructor-led PHR training.

The second stage of the structural analysis would therefore be based on a sample that was still

further restricted. For these reasons, PLS was selected as the analysis technique. For purposes

of applying partial least squares to the model hypothesized in this study, SmartPLS 2.0 (Ringle,

Wende, & Will, 2005) was selected. It is one of a number of tools that can be used to perform

partial least squares analysis (Temme, Kreis, & Hildebrandt, 2010).

   The process for performing partial least squares analysis was based on the approach

recommended by Chin (2010) and Gotz, Liehr-Gobbers, and Krafft (2010). This researcher

recommended a process that involved examination of the measurement or "outer" model,

followed by evaluation of the path or "inner" model. The steps in the analysis of this study were

as follows:

1. Evaluated the measurement (outer) model for reliability and validity. Reliability was tested for latent variables. First, item reliability was tested to determine the loading of the individual measurement items on their latent variables. Next, the latent variable reliability was tested using Cronbach's coefficient alpha (SPSS, SmartPLS) and composite reliability (SmartPLS). Convergent validity was then tested using average variance extracted (AVE) as the test statistic and confirming a value greater than 0.5 (Chin, 1998; Fornell & Larcker, 1981). Finally, discriminant validity for these measures was also tested by evaluating average variance extracted and comparing the square root of its value to the latent variable's intercorrelations with other latent variables (Fornell & Larcker, 1981). These validity tests were performed using output available from SmartPLS.

2. Evaluated the structural (inner) base model for all hypotheses except for those that included the instructor-related factors of feedback and instructor expectations. In this step, the path (inner) model was evaluated for all respondents. This included both an analysis of path coefficient weights as well as testing significance of those path coefficients using t values. Those t values were derived using SmartPLS's bootstrapping feature with 500 sample values selected. These were used to evaluate statistical significance, where significance was defined for the directional hypotheses as $p < .05$. In addition, R^2 values were evaluated for all endogenous variables.

3. Evaluated the path (inner) model for Ho1 through Ho6 with data from those respondents that indicated participation in instructor-led training. In this step, the instructional factors of instructor expectations and feedback were added to the

SmartPLS model. This analysis step employed the same tests of path coefficients, t values (and associated p values), and R^2 as described in the second analysis step.

4. Determined whether time on task and perceived usefulness mediated the effects of task-technology fit on PHR training self-efficacy (Ho7) and whether self-set goals mediated the effects of instructor expectations on PHR certification exam score (Ho8). For Ho8, significance of the mediation effect was determined using the approach described by Hubona (2011) and used by Helm, Eggert, and Garnefeld (2010). Also, the multiple mediation analysis approach described by Preacher and Hayes (2008) was used in the assessment of Ho7. Output from this step included Sobel test results, significance statistics, variance accounted for (VAF) and R^2 values.

Summary

This chapter addressed the research design, population, sample, instrumentation, pilot study, data collection, and data analysis procedures used to answer the research questions discussed in Chapter 1.

CHAPTER 4

FINDINGS

Overview

The purpose of this study was to investigate the relationship between two independent technology variables, five independent training variables, three motivation variables, and the dependent variable of Professional in Human Resources or PHR® certification (Human Resource Certification Institute, Alexandria, VA, www.hrci.org) outcome (see Figure 1). In the course of the study, eight research questions were evaluated. The first research question tested the relationship between task-technology fit and the factors of time spent on practice exam questions, time on task and perceived usefulness, and the relationship between perceived usefulness and PHR training self-efficacy. The second research question evaluated the relationship between the factors of time spent on practice exam questions, time on task, and feedback and the variable PHR training self-efficacy. The third research question tested the relationship between the factors of type of training, instructor expectations and feedback and the variable self-set goals. The fourth research question investigated the relationship between PHR training self-efficacy and measures of self-set goals and PHR exam self-efficacy. The fifth research question evaluated the relationship between the variables of PHR exam self-efficacy and self-set goals and the measure PHR certification outcome. The sixth research question tested the relationship between the factors of time spent on practice exam questions and type of training and the measure PHR certification outcome. Next, the seventh research question investigated whether time on task and perceived usefulness mediated the relationship between task-technology fit and PHR training self-efficacy. Finally, the eighth research question tested

51

whether self-set goals mediated the relationship between instructor expectations and self-set goals.

This chapter reports the findings of the study. The Data Assessment and Descriptive Statistics section outlines results of sampling, including sample size, and reports descriptive statistics, including the mean, standard deviation, normality, and kurtosis of the observed variables, including the latent variable indicators. The Assessment of the Measurement Model, Analysis of the Structural Model, and Mediation Analysis sections discuss testing of the measurement model, including instrument reliability and validity, testing of the structural model, and tests of the mediation hypotheses. They also address the results of tests of the null hypotheses. Chapter 4 concludes with a Summary section.

<center>Data Assessment and Descriptive Statistics</center>

Sample Size

After completion of the survey, 201 total responses had been received from examinees in 25 states. These responses were examined to select those that were from PHR examinees, met the date range criteria, completed both examination training and the exam itself, and exhibited no other significant anomalies in the respondents' replies. Completeness of data was also evaluated to determine which respondents to include in the analysis and which responses should be excluded. Of the 201 responses received, 127 were from PHR examinees. The remainder had taken the SPHR® (72) or, in two cases, presumably the GPHR® (Human Resource Certification Institute, Alexandria, VA, www.hrci.org). Of the 127 PHR responses, 9 were eliminated because of the elapsed time since their PHR exam preparation training had occurred (earlier than 1-2 mos. prior to May exam or 2-3 mos. prior to June exam). This resulted in 118 candidate PHR responses.

Of the 118 PHR responses that remained, 8 more were eliminated because of an excessive amount of missing data, which was largely attributable to respondents' stopping the survey prior to its completion. Three other responses with missing data were retained due to their exhibiting conditions that were not as well represented in the remainder of the data (i.e., high or low PHR exam scores or goals). This resulted in a sample size of 110 PHR responses, which was much larger than the sample size of 84 determined to have the desired power for performing the analysis using PLS path modeling. Examination of the data indicated that of those 110 responses, 89 were from respondents who had participated in instructor-led training and thus had responded to the two scales regarding instructor expectations and feedback. This was again above the desired threshold sample size of 84 discussed earlier. Thus, it was determined that these sample sizes were sufficient to proceed with PLS path modeling for both the initial model and the second model with two additional instructional factors, as planned in the study. Other activities that occurred during this phase included reverse coding for items requiring this step due to the phrasing of the survey item, addressing three response errors in the exam self-efficacy scale by treating them as missing data, and the replacement of any individual null items with an appropriate value so that they would be recognized as such by SmartPLS.

Descriptive Statistics and Common Method Bias

Table 1 shows relevant statistics from SPSS® (International Business Systems, Armonk, NY, www.ibm.com) for the constructs in the study using the unstandardized latent variable scores computed by SmartPLS as the source of input. Unlike covariance based SEM, PLS is not sensitive to violations of an assumption of normality. Nonetheless, univariate normality was analyzed using these descriptive statistical data. West, Finch, and Curran (1995) proposed that the cutoffs for univariate normality are skewness of less than 2 and kurtosis of less than 7.

Examination of the data in Table 1 showed that the measures analyzed in the study tended to be negatively skewed and leptokurtic, but all of the variables met the criteria for univariate normality.

Table 1

Descriptive Statistics for Study Variables

Measure	Mean	Std dev	Skewness	Kurtosis
Exam score	5.19	1.951	.066	-.524
Perceived usefulness	5.45	1.088	-1.302	3.668
PHR Exam Self-efficacy	30.54	9.819	.083	-.242
PHR Training Self-efficacy	4.74	1.271	-.402	-.066
Self-set goal	4.74	1.417	.320	-.325
Task-technology fit				
1. Work compatibility	5.51	1.149	-1.613	3.897
2. Ease of use	5.80	1.152	-1.666	4.193
3. Ease of learning	5.74	1.143	-1.576	3.978
4. Information quality	5.65	1.120	-1.555	4.191
Time on task	4.37	.669	-1.161	1.386
Time spent on exam quest	3.08	2.465	1.811	2.839
Type of training	4.89	1.149	-.945	-.959
Feedback*	3.07	1.098	-.175	-.770
Instructor expectations*	3.22	.858	-.451	.526

*From instructor-led training dataset with 89 responses (all others from 110 dataset).

For self-set goals, the approach was adapted from work done by Chen et al. (2000), where the researchers combined two test-grade goal items into a single self-set goal measure. Since "no goal" was an option in either case, the approach used in creating a composite goal

response was to scale the responses and then average them if both goals were provided, take the single goal response if only one was given, and treat the response as missing data if no value was given for either goal. Cronbach's alpha for the subset with both goal items was 0.63, but reconciliation of two incongruences yielded an acceptable .75 alpha. This composite self-set goal measure was then used in the subsequent analyses, following the approach used by Chen et al.

To test for common method bias, both Harman's single-factor test and partialling out of a social desirability measure were employed (Podsakoff et al., 2003). For Harman's single-factor test, factor analysis was used to determine if one factor explained the majority of the measurement items in either the initial model or the second instructor-led model. No such single factor was identified in either case. Next, following the SmartPLS modeling approach described by Elbashir, Collier, and Sutton (2011), a social desirability scale developed by Hays et al. (1989) was added to both models as a control variable on the endogenous variables. The correlation results from these models were reviewed and compared to the two zero-order models. In both cases, possible significant differences in the correlations were identified and then analyzed using the method described by Olkin and Finn (1995), as recommended by Podsakoff et al. To perform this analysis, SPSS syntax, provided by IBM (2010), was used to compare the zero-order and partial correlations. The results were that no significant differences were found between the zero-order correlations and the correlations with social desirability partialled out. Thus it was concluded that the original results were not affected by common method bias attributable to either a single factor or specifically to a social desirability factor.

Assessment of the Measurement Model

During the data screening, SPSS was initially used to test for reliability of the latent variable scales and compared with the values reported by the original researchers whose work was the source for the instrument. The test statistic used in this step was coefficient alpha. The responses were then loaded into SmartPLS 2.0 (Ringle et al., 2005), and the outer (measurement) model was evaluated using an approach similar to the one suggested by Chin (2010). Indicator loadings first were examined for reliability, and any indicators that did not exhibit reliability were removed. SmartPLS was then used to evaluate both composite reliability and coefficient alpha for the final measurement model. Both convergent and discriminant validity were evaluated as well for all of the latent variable items that were retained in SmartPLS. For both types of validity, an approach described by Fornell and Larcker (1981) was employed. Each of these tests employed average variance extracted (AVE) as the key metric of comparison. This information is also available in the analysis reports produced by SmartPLS. AVE above a value of 0.5 was considered to indicate good convergent validity, and discriminant validity was tested by confirming that the square root of the AVE measure for each latent variables was larger than its correlations with other variables (Chin, 1998; Fornell & Larcker, 1981).

Reliability

To test for internal consistency, coefficient alpha was first calculated in SPSS for all latent variables except feedback and instructor expectations. For those two variables, the subset of respondents who indicated that they participated in instructor-led training was utilized in the analysis. The results were then compared to the coefficient alphas reported by the researchers in the original research from which the instruments were derived and found to be acceptable. Next, indicator reliability was assessed for the latent variable indicators using the approach

recommended by Chin (1998). Several indicators were found to have loadings below the acceptable threshold value of 0.7 discussed by Gotz et al. (2010) and were removed. The reason for the 0.7 threshold is that, as Chin noted, in that case the indicator shares "more variance with the component score than with error variance" (p. 325). Following this approach, items that exhibited poor indicator reliability in either of the two samples were removed from the following scales: PHR training self-efficacy, instructor expectations, feedback, and time on task. This resulted in scales with four, five, four, and two indicators, respectively, achieving acceptable levels of indicator reliability and resulting in acceptable outcomes in the convergent validity measure of average variance extracted and the instrument reliability measures of composite reliability and coefficient alpha. In the case of time on task, items yielding a coefficient alpha value of .68 were selected given the overall result of statistically significant variable loadings, acceptable AVE and composite reliability measures, and consistent results between the two samples (the second alpha value was .71). This combined set of results was not found to be possible with other combinations of items from the instrument. The removal of measurement items represents a potential threat to nomological validity. Thus, while making the changes addressed the indicator reliability, and as a result, improved the construct validity aspects of convergent and discriminant validity, it potentially affects the nomological dimension of contruct validity. The resulting AVE, composite reliability, and coefficient alpha values from SmartPLS are shown in Table 2.

Table 2

Reliability and AVE Statistics From SmartPLS

	AVE	Composite reliability	Cronbach's alpha
Task-technology Fit			
- Work Compatibility	.694	.871	.773
- Ease of Use	.902	.964	.945
- Ease of Learning	.869	.952	.924
- Information Quality	.848	.944	.910
Perceived Usefulness	.837	.939	.903
Time on Task	.771	.871	.706
PHR Training Self-efficacy	.737	.917	.878
Instructor Expectations	.760	.941	.927
Feedback	.728	.914	.773

From instructor-led training dataset with 89 responses.

Convergent Validity

The formative construct task-technology fit was modeled in SmartPLS 2.0 using a second order molar approach as described by Chin (2010). This follows the approach used by Staples and Seddon (2004), who modeled it this way in their original research. It was also appropriate given that the construct met the criteria recommended by Chin that it be used when there are an equal number of indicators for each first-order construct that comprises the second-order construct. Convergent validity was then evaluated for each of the four factors that comprise task-technology fit (work compatibility, ease of use, ease of learning, and information quality). In addition, for the three other latent variables in the base model (perceived usefulness, time on task, and training self-efficacy), as well as the two included for respondents who participated in instructor-led training (instructor expectations and feedback), convergent validity was also evaluated.

The method used was based on the approach described by Fornell and Larcker (1981). In this method, AVE is calculated as the sum of the squared correlations between each indicator and the latent variable. These results, as computed by SmartPLS, are shown in Table 2. In each case, the AVE was above the 0.5 value the authors described as indicating good convergent validity (Chin, 1998; Fornell & Larcker, 1981). This demonstrated that the variance captured by these latent variable constructs was larger than the variance due to measurement error.

Similarly, convergent validity was evaluated for the two variables that were combined to form the self-set goal construct. The same general approach was followed, but in this case, the formula utilized SPSS output, as described by Ping (2002). Although the two observed variables were part of the measurement model, they were consolidated into one composite variable in the structural model, similar to the analysis approach used in the original study by Chen et al. (2000). The result of this evaluation was an average variance extracted of 0.61. This result was again above the 0.5 threshold value, indicating that the variance captured from the self-set goal construct was greater than the variance due to measurement error. These results, together with the reliability measures reported earlier, demonstrated that the latent variables in the model as well as the composite variable self-set goals were all internally consistent and exhibited convergent validity.

Discriminant Validity

For purposes of determining discriminant validity, Fornell and Larcker (1981) again provided a method that used average variance extracted. In this case, the method called for an evaluation of the AVE measure of a variable as compared to its squared correlations with other variables in the model. For purposes of performing this analysis, correlations were obtained from SmartPLS since, for the latent variables, correlations could not be evaluated with SPSS.

The results of this discriminant validity analysis are shown in Tables 3 and 4, using values obtained from SmartPLS with the instructor-led training dataset to allow evaluation of the complete set of variables. Following the method proposed by Fornell and Larker and using the presentation approach described by Sosik et al. (2009), discriminant validity was demonstrated for all constructs. This was indicated by the fact that the square root of the AVE measure for each latent variables (bolded and on the diagonal) was larger than its correlations with the other latent and manifest variables.

Table 3

Discriminant Validity Analysis (1 of 2)

	EL*	EU*	ExamSE	FeedBk	IQ*	InstExp	PU
Ease of Learning*	**.932**						
Ease of Use*	.859	**.950**					
Exam Self-efficacy	.120	.098	**1.000**				
Feedback	-.112	-.036	.086	**.853**			
Information Quality*	.849	.809	.148	-.029	**.921**		
Instructor Expectations	-.118	-.130	.048	.470	-.075	**.872**	
Perceived Usefulness	.758	.663	.226	-.003	.644	.117	**.915**
Self-set Goals	.043	.095	.315	.247	.085	.141	.104
Score	.011	.001	.270	-.118	-.076	.043	.034
Time Spent on Questions	.004	.026	-.010	.175	.042	-.164	-.024
Time on Task	.158	.208	.096	.285	.201	-.014	.152
Training Self-efficacy	.199	.190	.472	.161	.194	.094	.263
Type of Training	.045	.029	-.058	-.138	.048	.107	.149
Work Compatibility*	.721	.691	.127	.171	.679	.152	.720

*Task-technology Fit component factor.

Table 4

Discriminant Validity Analysis (2 of 2)

	SSG	Score	ToQ	ToT	TrnSE	TypTrn	WC*
Ease of Learning*							
Ease of Use*							
Exam Self-efficacy							
Feedback							
Information Quality*							
Instructor Expectations							
Perceived Usefulness							
Self-set Goals	**.782**						
Score	.287	**1.000**					
Time Spent on Questions	.037	-.267	**1.000**				
Time on Task	.112	-.086	.284	**.878**			
Training Self-efficacy	.299	-.004	.038	.355	**.858**		
Type of Training	-.017	-.061	-.055	-.100	-.077	**1.000**	
Work Compatibility*	.097	-.073	.129	.412	.312	.054	**.833**

*Task-technology Fit component factor.

Analysis of the Structural Model

After completion of the data assessment and measurement model assessment, analysis of the structural or inner model was performed using SmartPLS 2.0 (Ringle et al., 2005). The first model was analyzed using the entire 110 respondent dataset, and it included all of the variables except for feedback and instructor expectations. These two variables were excluded because this information was not requested from those who had not participated in instructor-led training. Then, analysis was performed on a second model that included all of the variables and utilized the dataset of 89 respondents who had participated in instructor-led training.

Both analysis steps included assessment of path coefficient weights as well as testing the significance of those path coefficients using *t* values (Chin, 2010). The *t* values were derived

using SmartPLS's bootstrapping feature with 500 sample values selected. These test statistics were used to evaluate statistical significance, where significance was defined for the directional hypotheses as $p < .05$ ($t > 1.658$). In addition, R^2 values were evaluated for all endogenous variables. Next, the research questions were evaluated using the results of the two analyses. For the final two research questions, additional steps were taken, including the calculation of Sobel statistics, bootstrap confidence intervals, and R^2, to assess the presence of mediation and its statistical significance.

Base and Instructor-led Model Analyses

First, PLS analysis was performed on the base model using SmartPLS 2.0. The resulting path coefficients and significance levels are shown by hypothesis in Table 5 (columns 2 & 3). In addition, the path coefficients, along with the resulting R^2 values for the endogenous variables, are shown graphically in Figure 4, using an approach similar to Chin (1998). Finally, the R^2 values computed by SmartPLS 2.0 are shown in Table 6 for the endogenous variables.

Next, these same steps were performed using the instructor-led model. This model included the additional variables of instructor expectations and feedback which were obtained only from respondents who indicated that they had participated in instructor-led training. The resulting path coefficients and significance levels from this second analysis step are shown in columns 4 and 5 of Table 5. Also, the path coefficients, along with the resulting R^2 values for the endogenous variables, are shown graphically in Figure 5. Finally, the R^2 values computed using SmartPLS 2.0 are shown in Table 6 for the endogenous variables.

Table 5

Model Path Coefficients and p Values

| Hypothesis and Path	Base model		Instructor-led model	
	Path coefficient	*p* value	Path coefficient	*p* value
H1a: Task-technology fit to Time spent on exam questions	.116	.071	.050	.303
H1b: Task-technology fit to Time on task	.248	.003	.261	.003
H1c: Task-technology fit to Perceived usefulness	.826	<.0001	.769	<.0001
H1d: Perceived usefulness to PHR training self-efficacy	.286	.002	.214	.034
H2a: Time spent on exam questions to PHR training self-efficacy	.014	.448	-.061	.286
H2b: Time on task to PHR training self-efficacy	.260	.005	.317	.004
H2c: Feedback to PHR training self-efficacy			.029	.217
H3a: Type of training to Self-set goals	-.056	.238	.029	.361
H3b: Instructor expectations to Self-set goals			.018	.444
H3c: Feedback to Self-set goals			.200	.044
H4a: PHR training self-efficacy to Self-set goals	.241	.003	.268	.002
H4b: PHR training self-efficacy to PHR exam self-efficacy	.524	<.0001	.472	<.0001
H5a: PHR exam self-efficacy to PHR certification exam score	.226	.005	.189	.034
H5b: Self-set goals to PHR certification exam score	.198	.007	.237	.006
H6a: Time spent on exam questions to PHR certification exam score	-.186	.019	-.278	.003
H6b: Type of training to PHR certification exam score	.015	.440	-.061	.246
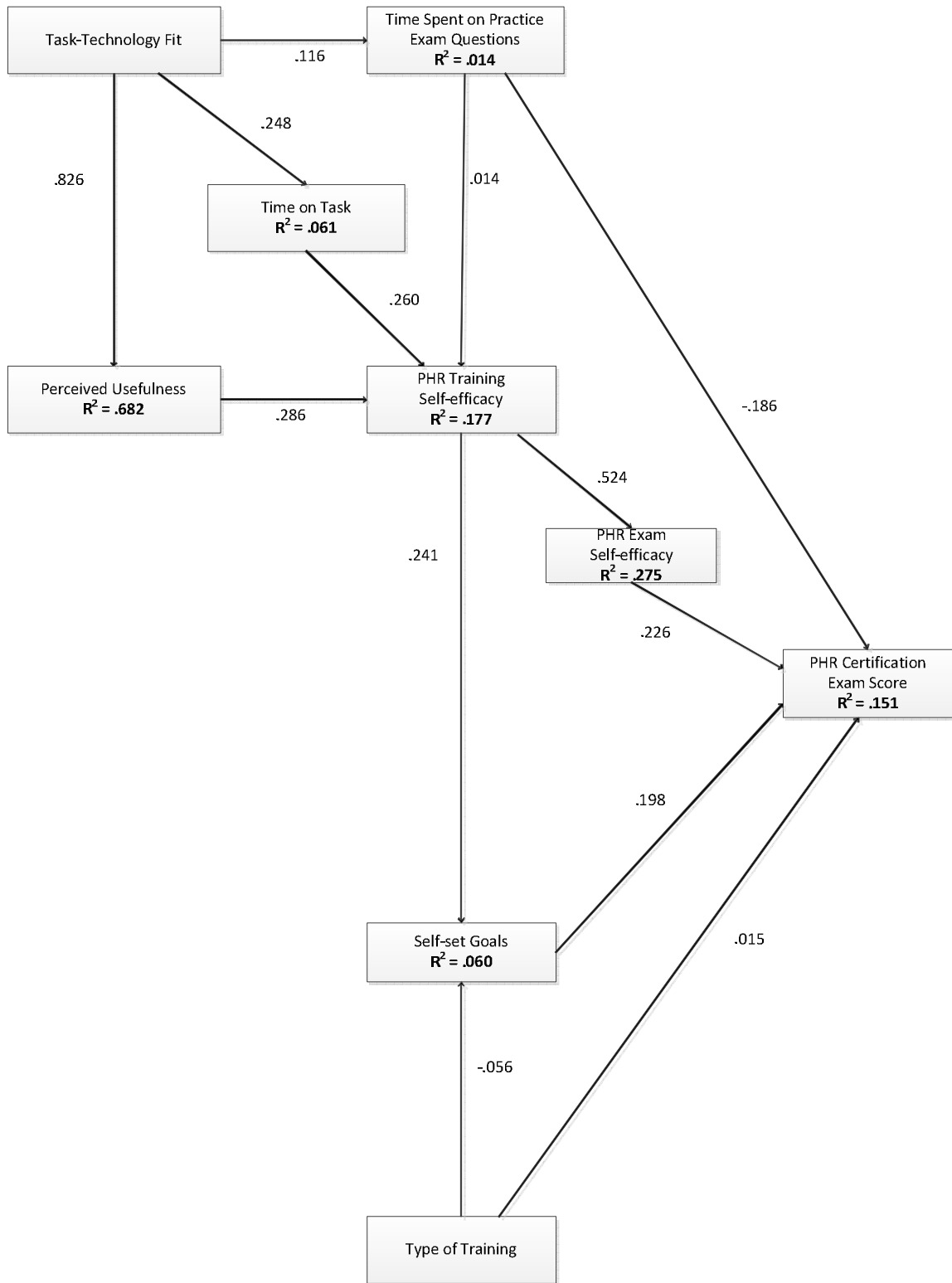
Figure 4. Base model.

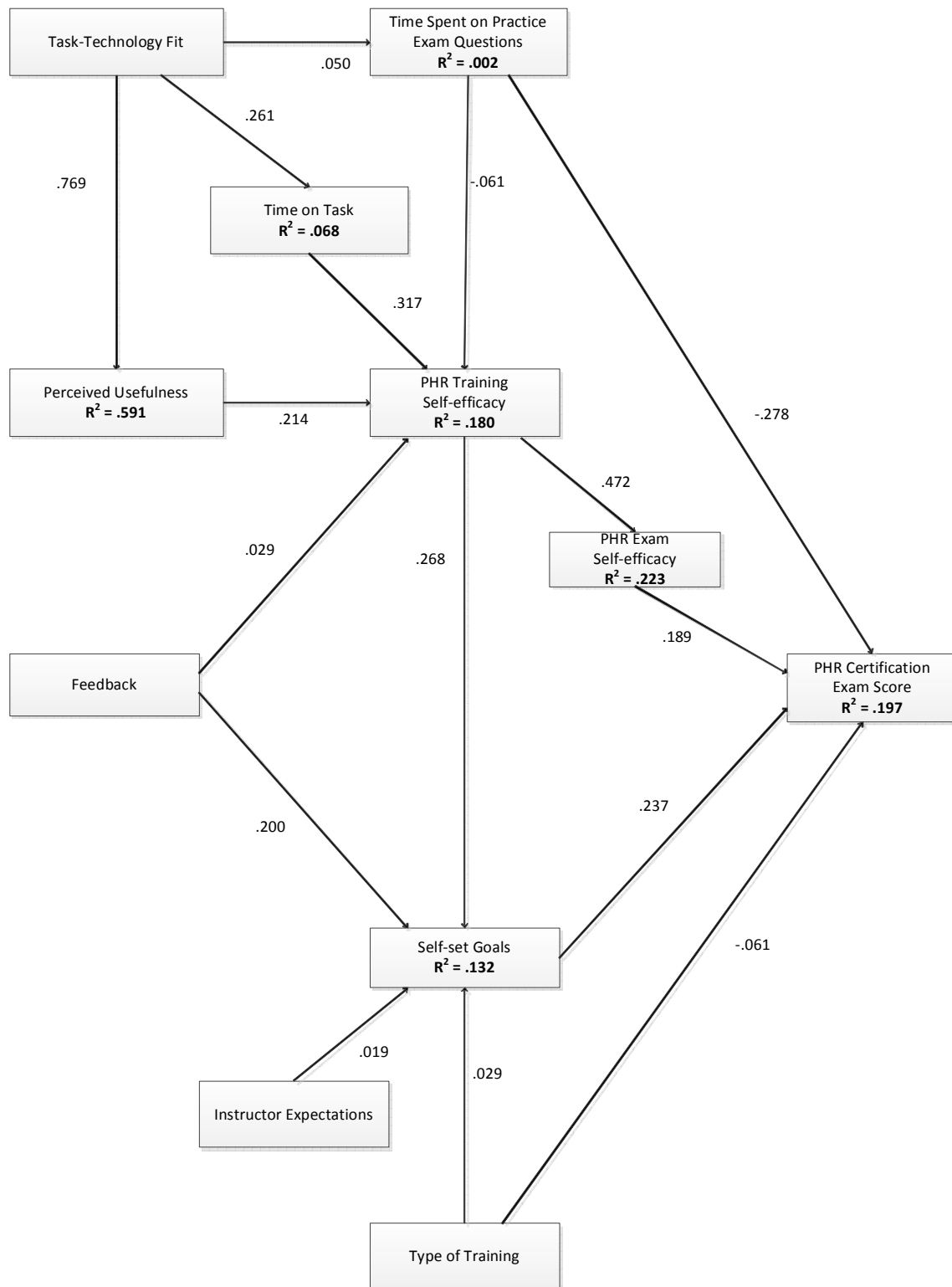*Figure 5. Instructor-led training model.*

Table 6

R^2 Values for Endogenous Variables

	Base model R^2	Instructor-led training model R^2
Perceived Usefulness	.682	.591
Time on Task	.061	.068
Time Spent on Exam Questions	.014	.002
PHR Training Self-efficacy	.177	.180
PHR Exam Self-efficacy	.275	.223
Self-set Goals	.060	.132
PHR Certification Exam Score	.151	.197

Evaluation of Research Questions 1-6

Next, the path coefficients and significance levels were interpreted to determine whether they demonstrated support for the two models' underlying hypotheses. This analysis was performed for both the base model and the instructor-led training model. For both of these evaluations, Cohen's (1988) general guideline for product-moment correlation and multiple linear regression was also used to assess the effect sizes. Thus, a standardized path coefficient with absolute value equal to .10 was considered a "small" effect; .30 was considered a "medium" effect; and .50 was determined to be a "large" effect (Cohen, 1988). These effect sizes are consistent with values used in covariance-based SEM as well (Kline, 2005). The analyses of the hypotheses and associated effect sizes analyzed in this step are shown in Table 7 for both the base model and the instructor-led training model. These results demonstrated support for a number of the hypotheses as shown in Table 7. In addition, several effects exhibited a change in their effect sizes, and several R^2 values changed when the instructor-led training condition was modeled. The interpretation and implications of these various results are addressed in Chapter 5.

Table 7

Hypotheses Supported and Effect Sizes

Hypothesis and path	Base model		Instructor-led model	
	Supported?	Effect size	Supported?	Effect size
H1a: Task-technology fit to Time spent on exam questions	No	Small	No	
H1b: Task-technology fit to Time on task	Yes	Sm/Med	Yes	Sm/Med
H1c: Task-technology fit to Perceived usefulness	Yes	Large	Yes	Large
H1d: Perceived usefulness to PHR training self-efficacy	Yes	Medium	Yes	Sm/Med
H2a: Time spent on exam questions to PHR training self-efficacy	No		No	
H2b: Time on task to PHR training self-efficacy	Yes	Sm/Med	Yes	Medium
H2c: Feedback to PHR training self-efficacy			No	
H3a: Type of training to Self-set goals	No		No	
H3b: Instructor expectations to Self-set goals			No	
H3c: Feedback to Self-set goals			Yes	Sm/Med
H4a: PHR training self-efficacy to Self-set goals	Yes	Sm/Med	Yes	Medium
H4b: PHR training self-efficacy to PHR exam self-efficacy	Yes	Large	Yes	Large
H5a: PHR exam self-efficacy to PHR certification exam score	Yes	Sm/Med	Yes	Sm/Med
H5b: Self-set goals to PHR certification exam score	Yes	Sm/Med	Yes	Sm/Med
H6a: Time spent on exam questions to PHR certification exam score	No*	Sm/Med	No*	Sm/Med
H6b: Type of training to PHR certification exam score	No		No	

*Not supported because of the negative relationship vs. hypothesized relationship direction.

Mediation Analysis

The final step in the analysis process consisted of assessing the mediation effects hypothesized in H7 and H8. The conditions defined by Baron and Kenny (1986) as necessary for mediation were tested. These involved confirmation of the existence of relationships between the independent and dependent variable (path c), independent to proposed mediator (path a), and mediator to dependent variable (path b). Then the Sobel test described by MacKinnon, Lockwood, Hoffman, West, and Sheets (2002) was used to evaluate the significance of the mediation relationships. That statistic and the associated p value were then reported. For research question 8, the instructor-led model dataset was used. The analysis for H8 was performed using information from SmartPLS and an online Sobel Test calculator available from Preacher (2011). Also, the variance accounted for (VAF) was calculated as an estimate of the magnitude of the indirect mediation effect. The approach taken was consistent with the discussion by Helm et al. (2010) and Hubona (2011). This information is reported in Table 8.

For both parts of research question 7, normal theory tests, bootstrapping, and an R^2 effect size calculation were performed using an SPSS script developed by Hayes (2009) for analyzing multiple mediation models. This SPSS script is called INDIRECT, and Hayes developed it based on the multiple mediation model analysis approach described by Preacher and Hayes (2008). These researchers made the point that the Sobel test works best with large sample or effect sizes and that bootstrapping is their primary recommendation. For this analysis, latent variable scores calculated by SmartPLS for the four variables involved (IV, DV, and two MVs) were extracted from the instructor-led model results and loaded into a file for the purpose of running the script in SPSS. Hayes's script calculated the Sobel test statistic and p value for both indirect paths as well as the total indirect effect, bootstrap confidence intervals for the indirect

effects, an overall R^2 value for the DV model, and its associated p value. In addition, the variance accounted for (VAF) was again calculated using path coefficients from individual SmartPLS models developed to test H7a and H7b. These VAF values served as estimates of the magnitude of the indirect mediation effects (Hubona, 2011). The results of this analysis are shown in Table 8.

Table 8

Mediation Model Statistics and Significance

Statistic	H7a: Perceived usefulness	H7b: Time on task	H7: Total indirect effect	H8: Self-set goals
Variance accounted for	76.2%	39.4%		87.5%
Sobel test	1.87	1.98	2.42	1.07
p value	.062	.047	.016	.287
Normal theory effect	.269	.108	.378	
Bootstrap confidence intervals	(-.022, .566)	(.012, .240)	(.040, .754)	
R^2 for DV model			.171	
p value for DV model			.001	

Evaluation of Mediation Research Questions 7 and 8

Using the Preacher and Hayes (2008) approach and SPSS script for testing multiple mediation, Ho7 was rejected since the p value for the total indirect effect (.016) was statistically significant, as supported by its bootstrap confidence interval (.040, .754) which did not include zero. The p value of the Sobel test for perceived usefulness (.062) was slightly above the significance level of $p < 0.05$, and the confidence interval for the bootstrap results (-.022, .566)

did include zero. Thus, the mediation effect for this indirect path was not found to be statistically significant. However, time on task was found to be a significant mediator of the relationship between task-technology fit and PHR training self-efficacy based on the statistically significant Sobel test result (p value = .047) and bootstrap confidence interval (.012, .240). Overall, the fact that the total indirect effect size is larger than the effect size for the time on task indirect path alone suggested that time on task and perceived usefulness together mediate the relationship between task-technology fit and PHR training self-efficacy. However, despite the larger effect size and VAF for the perceived usefulness indirect path, time on task was the only indirect path in this multiple mediation relationship that individually exhibited statistical significance. Finally, Ho8 was not rejected since the Sobel test result was clearly not significant. The VAF statistic for the self-set goals mediation relationship was large, but it was based on small effect sizes. Each of these results is discussed further in Chapter 5.

Summary

Chapter 4 reported the study's findings. The Descriptive Statistics section investigated the mean, standard deviation, normality, and kurtosis of the observed variables, including the latent variable indicators. The Assessment of the Measurement (Outer) Model section outlined results of analysis of instrument reliability, item reliability, convergent validity, and discriminant validity. The Analysis of the Structural (Inner) Model section addressed the results of tests of the null hypotheses, including the mediation hypotheses. Chapter 5 provides a discussion of the study's findings and presents conclusions, recommendations, and implications.

CHAPTER 5

DISCUSSION

Overview

This chapter includes the following sections: Synthesis of Findings, Discussion, Recommendations for Future Research, Implications for Practice, and Summary. In the Synthesis of Findings section, the researcher provides a summary of the results of the hypotheses findings to answer the study's research questions. The Discussion section draws conclusions from those findings and considers the limitations of the study. The Recommendations for Future Research section discusses topics for future investigators to address. The Implications for Practice section discusses findings relevant to instructional designers developing Professional in Human Resources or PHR® certification (Human Resource Certification Institute, Alexandria, VA, www.hrci.org) training, as well as to examinees themselves. It also integrates the study's findings more broadly into the field of Training and Development. The Implications for Theory section addresses how the findings add to the body of knowledge in several domains, and finally, the chapter concludes with a Summary section.

Synthesis of Findings

The purpose of this study was to investigate the relationships between two technology variables (task-technology fit and perceived usefulness), five training variables (time on task, time spent on practice exam questions, feedback, instructor expectations, and type of training), three motivation variables (PHR training self-efficacy, PHR exam self-efficacy, and self-set goals) and the dependent variable PHR certification exam score (see Figure 1). It also tested whether three of these variables acted as mediators.

The first research question investigated the relationship between task-technology fit and the factors of time spent on practice exam questions, time on task, and perceived usefulness, and the relationship between perceived usefulness and PHR training self-efficacy. The next two research questions addressed whether or not there were relationships between the five instructional variables and the motivation variables PHR training self-efficacy and self-set goals, and PHR certification exam score. The fourth and fifth research questions investigated whether there were relationships between those same two motivation variables and between them and PHR exam self-efficacy and the dependent variable PHR certification exam score. The sixth research question addressed whether or not two of the instructional variables (time spent on practice exam questions and type of certification training) had a direct effect on PHR certification exam score. The seventh research question addressed whether perceived usefulness and time on task mediate the relationship between task-technology fit and PHR training self-efficacy. Finally, the eighth research question tested whether self-set goals mediate the relationship between instructor expectations and the dependent variable, PHR certification outcome. All eight research questions and their associated null hypotheses were evaluated using the partial least squares (PLS) analysis tool SmartPLS 2.0.

The path coefficients and significance levels for both the base model and the instructor-led training model are shown in Figures 4 and 5. As shown in those models, three of the four relationships evaluated in the first research question were found to be statistically significant. The effects of task-technology fit on time on task (.248, .261) and perceived usefulness (.826, .769), as well as the effects of perceived usefulness on PHR training self-efficacy (.286, .214), were all significant at the $p < .05$ level or greater. The one path in this research question not found to be significant was the effect of task-technology fit on time spent on practice exam

72

questions, although it reached a .10 significance level in the base model. For the second and third research questions, two statistically significant effects were found between the instructional factors and the motivation variables. First, there was a significant relationship between time on task and PHR training self-efficacy, with path coefficients of .260 and .317 in the two models (p < .01 for both). The relationship between feedback and self-set goals, with a path coefficient of .200, was also found to be statistically significant at the $p < .05$ level.

In the fourth research question, statistically significant relationships were found between PHR training self-efficacy and both self-set goals (.241, .268) and PHR exam self-efficacy (.524, .472). In both cases, the associated t values were significant at the $p < .01$ level. For the fifth research question, the path coefficients for the PHR exam self-efficacy to PHR certification exam score path (.226, .189) were found to be statistically significant at the $p < .01$ and .05 level, respectively. Also, for the self-set goals to PHR certification exam score path, the path coefficients for the two models (.198, .237) were both significant at the $p < .01$ level. The sixth research question investigated the direct effects of time spent on practice exam questions and type of training on PHR certification exam score. Only the path coefficients for the first relationship were found to be statistically significant (-.186, -.278; $p < .05, p < .01$). However, the negative direction associated with this relationship meant that the null hypothesis was nonetheless accepted.

Finally, the seventh and eighth research questions addressed mediation hypotheses. In the seventh research question, the roles of time on task and perceived usefulness in mediating the relationship between task-technology fit and PHR training self-efficacy were established. The time on task indirect path was found to be significant, with a p value of .047 while the perceived usefulness indirect path was not (p = .062). For the full multiple mediation relationship

73

addressed in research question 7, the statistically significant *p* value (.016) of the total indirect effect provided support for the multiple mediation hypothesis. These conclusions were supported by analyses of bootstrap confidence interval results as recommended by Preacher and Hayes (2008). The multiple mediation model is discussed in more detail in the next section. For research question 8, however, self-set goals were not found to mediate the relationship between instructor expectations and PHR certification exam score.

Discussion

The following section describes conclusions that were made based on the outcomes of the study. The limitations section addresses uncontrollable elements of the study. These uncontrollable items represent threats to validity that should be considered when looking at the generalizability of the study's results.

Conclusions From the Findings

In the first three research questions, four effects were found to be statistically significant. As hypothesized, task-technology fit was found to have an effect on both time on task and perceived usefulness. Then, in turn, perceived usefulness and time on task significantly affected PHR training self-efficacy. This supports the hypothesized model in which instructional and technology factors work through psychological factors, including motivation, to affect learning outcomes. In addition, the results directly support the findings of Staples and Seddon (2004) regarding the relationship between task-technology fit and perceived usefulness and as well, the findings of Cheung et al. (2003) regarding the relationship between perceived usefulness and learning self-efficacy. It could also be concluded that the factors that affect how well technology supports training can in turn affect training self-efficacy. This point is further examined in research question 7 and is an important link in establishing the relationship between task-

74

technology fit and ultimate outcomes. Finally, the lack of a significant effect on self-set goals from the Chickering and Gamson (1987) factor of instructor expectations is interesting as well. This may be related as much to the way this type of training is delivered as it is based on the potential for the factor to play a role in influencing motivational factors. Nonetheless, it does not support the findings of Klomegah (2007) who reported a significant correlation between assigned goals and self-set goals. Feedback, the third Chickering and Gamson factor examined in the study was again shown to play a significant role in goal setting. This supports prior research by Latham and Locke (1991) and as well, the idea that feedback is an important component of the instructional process. The lack of a significant relationship between feedback and PHR training self-efficacy is very interesting, however, since it does not support studies by a number of prior researchers (Bandura & Cervone, 1983; Latham & Locke, 1991; Schunk, 1990), who reported results showing that feedback affects self-efficacy. This is obviously a surprising outcome and one that suggests the need for follow-up evaluation.

The results of the fourth and fifth research questions demonstrated the usefulness of the motivation hub in predicting outcomes, and specifically, score outcomes for the PHR certification exam. PHR training self-efficacy was shown to affect PHR exam self-efficacy and self-set goals, and both PHR exam self-efficacy and self-set goals were found to impact PHR certification exam scores. These are not surprising outcomes; they support prior theory (Latham & Locke, 1991) and empirical research regarding academic outcomes (Chen et al., 2000; Phillips & Gully, 1997; Wood & Locke, 1987). The motivation hub plays an important role in this path model since it completes the linkage from the two technology and instructional factors through to exam score outcome.

The sixth research question was based on theory, but it also specifically considered empirical research in the domain of certification exams. Its investigation led to the identification of a factor that significantly affects PHR examination exam scores, but just as importantly, it suggested one that does not have an effect. First it suggested that time spent on practice exam questions was negatively related to exam outcomes. Discovery of this negative relationship meant that the directional null hypothesis was accepted. This significant negative path coefficient is a new finding that requires further investigation. As discussed later in the Recommendations for Future Research section, it is suggested that such research should investigate whether the relationship is nonlinear, and whether repeated use of practice exams by underprepared examinees is a factor. Equally important to consider is the lack of a significant finding regarding the effect of type of training on PHR certification exam scores. This outcome does not support the results of Grange et al. (2003), who found that there was a difference in Certified Financial Planner™ (CFP Board, Washington, DC, www.cfp.net) exam outcomes based on the type of preparation. Perhaps in the domain of PHR exams, it does not play the same role as elsewhere; however, it is an important finding to note given the ongoing discussions about the role of online versus classroom training activities.

Finally, two types of mediation models were addressed in the seventh and eighth research questions. In the seventh research question the indirect linkage from task-technology fit to the motivation hub was investigated. The combined roles of time on task and perceived usefulness in mediating the relationship between task-technology fit and PHR training self-efficacy were established. This is a multiple mediation relationship that combines a technology-related factor with an instructional factor. Only the time on task indirect path was shown to be individually significant; however, the total indirect relationship with both mediating variables was not only

76

statistically significant but also had a much larger effect size than the time on task indirect path alone. This is a new finding, and it suggests that the interplay of technology and instructional factors can be complex and requires further investigation. Surprisingly, however, analysis of research question 8 failed to lead to the conclusion that self-set goals mediated the relationship between instructor expectations and the dependent variable PHR certification exam score. Based on the earlier results, the self-set goals variable affects score outcomes, but it does not mediate the effect of instructor expectations on those outcomes.

Limitations

Several uncontrollable factors associated with the data evaluated in this study should be considered when evaluating the outcomes. First, response error and bias were not controlled, given that an online measurement instrument was used to collect the data. In addition, since factors such as PHR certification exam test scores were self-reported, the results were dependent on respondents being honest about the data they provided. Also, respondents were asked for these data after they had completed the PHR exam. Therefore, they may not have accurately recalled and reported their pretraining self-efficacy and self-set goals, although the time between those events was a variable that was analyzed and controlled. All of these represent threats to conclusion validity.

Also, measurement items were removed from several of the constructs to improve their indicator reliability, and as a result, their convergent and discriminant validity. The affected measures were reflective, and this provides some level of robustness. Nonetheless, the fact that it was necessary to make the changes to address the convergent and discriminant validity aspects of the constructs represents a threat to their nomological validity.

In addition, the research design of this study was not experimental, and it thus did not completely control for the effects of other variables. Also, the study included no longitudinal data or analysis. For these reasons, although mediation relationships are typically interpreted to imply causation, the ability to draw such strong conclusions is limited. That constraint represents a threat to the study's internal validity.

Finally, limitations in the sample population, along with the fact that random selection and assignment were not used, affect the ability to generalize beyond the study population. This represents a threat to the external validity of the study and its results. Each of these limitations could be addressed in future studies that modify the research design, expand the population, or gather data using direct methods. Further study is certainly warranted, and specific recommendations regarding such steps are made in the next section.

Recommendations for Future Research

While this study addressed a number of research questions, several areas are recommended for future researchers to consider, including extensions to the population, type of certification program, scope of outcome, demographics, and other personal factors that could be evaluated in such future studies. These recommendations include steps that would extend the results of the study and address some of the major threats to study validity, thereby improving its generalizability.

Experimental Design

The major threat to the study's internal validity was the fact that it was not an experimental design and included no longitudinal data. As a result, conclusions regarding causality are limited. Thus, one of the first recommendations to future researchers is that the study be replicated using an experimental design, including sampling throughout the study

period. For instance, the researcher might change the delivery of the training to include pre- and posttests along with a switching replications experimental design approach. Such a study would have much higher internal validity.

Population

The current study considered a specific population that was available through the sampling approach of contacting university-sponsored training programs as well as members of local Society for Human Resource Management chapters. As a result, the self-selection process led to a population in which the majority of respondents received instructor-led training. To improve the external validity of the study, and thus the generalizability of its results, further studies should be performed with populations that included a wider variety of training options. For instance, a potential research question might be whether those who take college-related classroom training, when compared to other populations, are more subject to the motivational effects demonstrated in this study.

Other Certification Programs

This study considered no certification programs other than the PHR exam. It is recommended that these same relationships be evaluated for other similar certification training activities. Certification exam programs such as the Certified Public Accountant and Certified Financial Planner have similarities including testing based on codified bodies of knowledge. The effect of the mechanisms in the motivation hub and their role in certification exam training programs should be consistent; however, that hypothesis should obviously be tested.

Skills Transfer

The scope of this study was narrowly focused on the effect of training on certification testing results. It was not designed to measure the impact of the training on skills transfer or career outcomes. Such an extension of the results is suggested, because no similar studies of the effects of training on certification program-related improvements in job skills or career outcomes were encountered during the review of the literature. Specifically, future research could compare the post-certification HR skills self-efficacy of trainees to their posttraining PHR exam self-efficacy to determine whether trainees retain a belief in their HR knowledge and ability is sustained or is only short-term and related to the exam itself. A study of such measures could also be extended longitudinally to consider relationships to factors such as career and salary progression.

Evaluation of Demographics and Personal Characteristics

A limited set of demographic data was obtained in the study, including years in the workforce, years in current position, industry, HR subfield, age range, gender, and level of formal education; however the demographics of the study population were not evaluated in this study. It is suggested that such an evaluation of study results by demographic category may indicate that differences in outcomes are influenced by such factors as years of professional experience. Also, prior educational background, academic history, grades, and other standardized test scores should be considered, given the role that ability has been found to play in other studies that have employed the motivational hub model to predict academic performance outcomes.

Evaluation of Other Factors

The training characteristics that were correlated with changes in self-efficacy and self-set goals were factors associated with the respondents' self-reported classroom, online, or self-

managed training preparations and the time they spent on training tasks and, specifically, on practice exam questions. While some data were gathered, no attempts were made to evaluate other informal or job-related learning factors such as supervisor support, financial work-related incentives, or mentoring. Also, no efforts were made to determine the effects of other training-related characteristics such as course content or design. Each of these factors could be included in subsequent studies and their impact on outcomes compared in terms of effect size. Such studies would help to clarify the role of goal setting and self-efficacy as compared to other training and job-related motivators.

Evaluation of the Effect of Time Spent on Practice Exam Questions

The factors examined in this research were evaluated using analysis techniques that assume a linear relationship between independent and dependent variables. However, some of the relationships may involve nonlinear effects. For instance, time spent on practice exam questions was found to have a negative effect on PHR certification exam scores. Perhaps this is a nonlinear relationship, and while studying practice exam questions may be useful to some degree, excessive use of them versus time spent understanding conceptual material is detrimental to success. Another possible explanation is that repeated use of practice exams is an approach employed by some individuals who are less well prepared for the exam in general and that these examinees achieve lower exam score outcomes. Whatever the explanation, obviously a more detailed analysis is suggested.

<div align="center">Implications for Practice</div>

The results from this study can be applied in the design and delivery of PHR certification programs, but they may be considered in other contexts as well. First, the results showed that task-technology fit can affect training self-efficacy through the mediating effect of time on task

<div align="center">81</div>

and an indirect path through perceived usefulness. The factor of training self-efficacy subsequently affected training outcomes through other mechanisms in the motivation hub. One conclusion from the result is that training technologies could be evaluated to determine trainee perceptions of task-technology fit and the resulting perceptions regarding its usefulness and that time spent in improving these factors could be valuable in affecting training results.

The study results also suggest specific interventions which can improve the motivational factors that affect PHR certification exam results. One such action is to increase the amount of time an examinee spends "on task" in the broader sense, preparing for the exam in a complete manner, as opposed to singularly focusing on taking practice exams and studying practice exam questions exclusively. Consistent with Bandura's (1997) concept of enactive mastery experiences, vicarious experiences, and verbal persuasion, the results indicate that time on task improves training self-efficacy and that this results in an increase in exam-specific self-efficacy and self-set goals. Through those mechanisms, test outcomes improve.

This result has implications for both curriculum development and classroom management. The more time PHR examinees spend on training tasks, the better the expected outcome. But, as noted earlier, solely focusing on practice exams has the opposite effect; however, this conclusion is based on the range of responses included in the study. Presumably, there is a point where the incremental benefits of increased time on task diminish. Similarly, practice exams are perhaps a good thing, which, when taken to an extreme, limit the amount of time required to build a broad base of knowledge necessary to be prepared for the breadth of questions that may be presented during the testing experience. Curriculum designers should keep these relationships in mind and balance the time spent on such exercises without going too far and ignoring the other learning domains. Instructors also must be creative in their classroom

82

management practices and training delivery to avoid the risk that such a training program

devolves into a tedious series of proctored practice exams. In addition, the benefits ascribed to

spaced practice (Donovan & Radosevich, 1999), as compared to massed practice, argue for

scheduling various forms of learning reinforcement, including practice exams, throughout the

duration of the training.

The type of training program selected made no difference; however, this outcome may

not be surprising given the availability and prevalence of PHR study programs that deliver

training through self-paced learning via study guides or Internet-based options. Nonetheless, the

results of this study did demonstrate beneficial effects from increases in the PHR certification

exam score goals that learners set for themselves. These self-set goals were significantly

affected by training self-efficacy, so its technology and time on task antecedents again come into

play. Of the three instructional factors tested for their impact on self-set goals, only feedback

had a path coefficient that was found to be statistically significant. This outcome further

reinforces the importance of providing frequent feedback during training no matter how the

training is delivered.

More generally, the study once again showed the applicability of the motivation hub

model in relating factors that can influence, and even predict, performance. As was discussed

earlier, there is no reason to think that this outcome is limited to the PHR certification exam, and

in fact, similar results have been demonstrated in academic domains. Although further studies

are warranted, in the meantime, instructional designers and instructors could decide to generalize

these results to other types of certification training programs and factor the roles of time on task,

time spent on practice exam questions, feedback, and task-technology fit into the overall design

of their programs. Thus, as instructional design decisions are considered, the mechanisms at

work in building self-efficacy and self-set goals could be taken into account and emphasized. Instructor behaviors and even interventions could be incorporated to achieve increases in these motivational factors, and through them, improve exam outcomes.

Implications for Theory

Aside from the specifics of the study, the results presented above have several theoretical implications. First, the result of the relationship between time on task and PHR training self-efficacy provides further empirical support for the importance of enactive mastery experiences, vicarious experiences, and verbal persuasion in the development of self-efficacy (Bandura, 1997). In this case, increased time spent on training activities predicted PHR training self-efficacy, which in turn affected two motivation factors that influence exam results. This is consistent with previous studies in the academic domain that have shown higher self-efficacy to be correlated with better exam outcomes. The significant path from feedback to self-set goals was also anticipated based on prior research. However, the lack of significance of the instructor expectations path was not expected, but was perhaps due to the domain of study. Also, the result regarding time spent on practice exam questions was significant but not in the expected direction. As discussed previously, this interesting result could imply a complex, perhaps nonlinear, relationship as well as interaction with other forms of exam preparation variables. Those alternatives can only be evaluated with further investigation.

The results regarding the mediation effects have several additional implications for theory. The study results suggest a multiple mediation model and a pathway for their indirect effects to ultimately affect score outcomes. As noted earlier, the role of time on task as a mediator and indirect role played by perceived usefulness in the relationship between task-technology fit and PHR training self-efficacy combine a technology-related factor and an

84

instructional factor in a multiple mediation model. Obviously, further research is warranted to better understand how these complex mechanisms operate. The fact that the self-set goals variable failed to mediate the relationships between instructor expectations and PHR certification exam score was not expected. Perhaps this was due to the specific domain of this study, but it again suggests the need for further study of this instructional variable.

As noted earlier, the conclusion that self-efficacy and self-set goals can affect test outcomes is not a new suggestion. The motivation hub construct has been extensively studied since its description by Locke et al. (1984). Many studies have in fact demonstrated such correlations in academic settings (Chen et al., 2000; Phillips & Gully, 1997; Wood & Locke, 1987). The relationship between the two forms of self-efficacy is also consistent with theory regarding the importance of considering domain-specificity when measuring the effect of self-efficacy. However, the study results demonstrate that in this subject domain, the variables participate in a complex model that includes mediation effects. These results ultimately add to our understanding of their effect and utility.

Summary

This study demonstrated the predictive ability of the motivation hub model on test score outcomes for the PHR certification exam. Specifically, PHR training self-efficacy was shown to affect PHR exam self-efficacy and self-set goals, which in turn impact PHR certification exam scores. It also demonstrated the role of task-technology fit and the indirect effects it has on PHR training self-efficacy through the mediation role played by time on task and indirect path through perceived usefulness. It also suggested that time spent on practice exam questions was negatively related to exam outcomes, although further research was suggested to determine

whether that relationship manifests itself linearly. Finally, it reinforced the role of feedback, specifically through its effect on self-set goals.

These results suggest that instructional designers developing training for PHR certification exams should consider how they could employ these results by improving the motivational as well as the content aspects of the programs. In addition, the results should be tested with other certification testing programs and populations, but it is anticipated that these theoretically supported relationships should translate into analogous results in other similar domains.

APPENDIX

FORMATIVE AND REFLECTIVE LATENT VARIABLES

Table A1

Formative Latent Variables – Items Analyzed

Construct	Type	Items
Task-technology fit	Formative	
1. Work compatibility	Reflective	
		Using the new system fit well with the way I like to study
		The system was compatible with all aspects of my studying
		I had ready access to the system when I needed it
2. Ease of use	Reflective	
		The system was easy to use
		The system was user friendly
		It was easy to get the system to do what I wanted it to do
3. Ease of learning	Reflective	
		The system was easy to learn
		It was easy for me to become more skillful at using the system
		New features were easy to learn
4. Information quality	Reflective	
		Do you think the output was presented in a useful format?
		Was the information accurate?
		Did the system provide up-to-date information?

Note. Adapted from "Testing the technology-to-performance chain model," by D. S. Staples and P. Seddon, 2004, *Journal of Organizational and End User Computing, 16*, p. 34. Copyright 2004 by Idea Group Inc. Reprinted with permission of the publisher.

Table A2

Reflective Latent Variables – Items Analyzed

Construct	Type	Items
Perceived Usefulness[a]	Reflective	
		The system enabled me to accomplish tasks more quickly
		Using the system increased my productivity
		Using the system made it easier to do my studying
PHR Training Self-efficacy[b]	Reflective	
		I believed I would receive an excellent score on the exam
		I was confident I could do an excellent job on exam questions like those in the assignments and tests for this course
		I expected to do well on the exam
		Considering the difficulty of the course, its delivery, and my skills, I thought I would do well on the exam
Time on task[c]	Reflective	
		I reviewed the material repeatedly
		In preparing for the exam, I completed:
		-None of the training exercises
		-A few of the training exercises
		-Some of the exercises
		-Most of the training exercises
		-All of the training exercises
Feedback[d]	Reflective	
		I conferred with my instructor if I was concerned about keeping up in the course
		I talked over feedback with my instructor as soon as possible if anything was not clear
		I did re-work and sought feedback from the

89

		instructor in doing so
		I sought feedback from my instructor about my work
Instructor expectations[e]	Reflective	For the average HR professional in your course, how difficult were your instructor's performance goals and expectations and what did they require?
		Extreme challenge ---- No challenge at all
		Enormous effort ---- Almost no effort
		An extreme degree of thought and problem solving skill ---- No thought or skill
		An enormous amount of persistence and tenacity ---- Very little persistence and tenacity
		Very high standards of performance ---- No standards of performance at all

Note. [a]Adapted from "Examining technology acceptance by school teachers: A longitudinal study," by P. J. H. Hu, T. H. K. Clark, and W. W. Ma, 2003, *Information & Management, 41*, p. 238. Copyright 2003 by Elsevier B. V. Reprinted with permission.
[b]Adapted from "The making of the motivated strategies for learning questionnaire," by T. G. Duncan and W.J. McKeachie, 2005, *Educational Psychologist, 40*, pp. 126-127. Copyright 2005 by Lawrence Erlbaum Associates, Inc. Reprinted with permission.
[c]Adapted from an instrument discussed in "Time-on-task mediates the conscientiousness-performance relationship," by M. D. Biderman, N. T. Nguyen, and J. Sebren, 2008, *Personality and Individual Differences, 44*, pp. 887-897. Instrument provided by and used with permission of author.
[d]Adapted from "Seven principles student inventory: An indicator of success?" by J. E. Oberst, 1995, Doctoral dissertation, p. 137. Copyright 1995 by J. E. Oberst. Reprinted with permission.
[e]Adapted from "Exploring the meaning and usefulness of measures of subjective goal difficulty," by C. Lee and P. Bobko, 1992, *Journal of Applied Social Psychology, 22*, p. 1428. Copyright 1992 by V. H. Winston & Son, Inc. Reprinted with permission.

Aguinis, H., Michaelis, S. E., & Jones, N. M. (2005). Demand for certified human resources professionals in Internet-based job announcements. *International Journal of Selection and Assessment, 13*(2), 160-171.

Alavi, M., & Leidner, D. E. (2001). Research commentary: Technology-mediated learning – A call for greater depth and breadth of research. *Information Systems Research, 12*(1), 1-10.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337), Greenwich, CT: Information Age Publishing.

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 45*(5), 1017-1028.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Biderman, M. D., Nguyen, N. T., & Sebren, J. (2008). Time-on-task mediates the conscientiousness-performance relationship. *Personality and Individual Differences, 44*, 887-897.

Brown, K. G. (2001). Using computers to deliver training: Which employees learn and why? *Personnel Psychology, 54*, 271-296.

Chen, G., Gully, S. M., Whiteman, J., & Kilcullen, R. N. (2000). Examination of relationships among trait-like individual differences, state-like individual differences, and learning performance. *Journal of Applied Psychology, 85*(6), 835-847.

Cheung, W., Li, E. Y., & Yee, L. W. (2003). Multimedia learning system and its effect on self-efficacy in database modeling and design: an exploratory study. *Computers & Education, 41*, 249-270.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin, 39*(7), 3-7.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295-336), Mahwah, NJ: Erlbaum.

Chin, W. W. (2010). How to write up and report PLS analyses. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods, and applications* (pp. 655-690), Berlin, Germany: Springer.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2*nd ed*).* New York: Taylor & Francis Group.

Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 83*(5), 678-707.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*, 36-60.

Dennen, V. P. (2005). From message posting to learning dialogues: Factors affecting learner participation in asynchronous discussion. *Distance Education, 26*(1), 127-148.

Dishaw, M. T., & Strong, D. M. (1999). Extending the technology acceptance model with task-technology fit constructs. *Information & Management, 36*, 9-21.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*(5), 795-805.

Duncan, T. G., & McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist, 40*(2), 117-128.

Dunn, W. M., & Hall, T. W. (1984). An empirical analysis of the relationships between CPA examination candidate attributes and candidate performance. *The Accounting Review, 59*(4), 674-689.

Eccles, J. S. (2005). Subjective task value and the Eccles et al. model of achievement-related choices. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105-121), New York: The Guilford Press.

Elbashir, M. Z., Collier, P. A., & Sutton, S. G. (2011). The role of organizational absorptive capacity in strategic use of business intelligence to support integrated management control systems. *The Accounting Review, 86*(1), 155-184.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analysis. *Behavior Research Methods, 41*(4), 1149-1160.

Fertig, J. (2011). Evaluating that piece of paper: The effect of motivation and certification status on occupational commitment and job competence. *Journal of Leadership & Organizational Studies, 18*(1), 118-126.

Fertig, J., Zeitz, G., & Blau, G. (2009). Building internal motivation for worker competency certifications: A critique and proposal. *Human Resource Development Review, 8*(2), 197-222.

Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology, 28*(2), 161-186.

Fletcher, K. M. (2005). Self-efficacy as an evaluation measure for programs in support of online learning literacies for undergraduates. *The Internet and Higher Education, 8*(4), 307-322.

Forman, D. C., & Cohen, D. J. (1999). The SHRM learning system – A brief history. *Human Resource Management, 38*(2), 155-160.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology. *Journal of Counseling Psychology, 51*(1), 115-134.

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly, 19*(2), 213-236.

Gotz, O., Liehr-Gobbers, K., & Krafft, M. (2010). Evaluation of structural equation models using the partial least squares (PLS) approach. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods, and applications* (pp. 655-690), Berlin, Germany: Springer.

Grange, E. V., Hampton, V. L., Cutler, R., Langdon, T. P., & Ryan M. T. (2003). Factors associated with success on the CFP certification examination. *Financial Services Review, 12*(2), 95-114.

Grant, C. T., Ciccotello, C. S., & Dickie, M. (2002). Barriers to professional entry: How

    effective is the 150-hour rule? *Journal of Accounting and Public Policy, 21*(1), 71-93.

Harrington, J., Davis, J., & Harrington, S. A. (1992). The use of PPST math scores to improve

    math skills for teacher education students. *Education, 112*(4), 523-528.

Hayes, A. F. (2009). INDIRECT SPSS macro (version 2.0). Retrieved from

    http://www.afhayes.com/spss-sas-and-mplus-macros-and-code.html

Hays, R. D., Hayashi, T., & Stewart, A. L. (1989). A five-item measure of socially desirable

    response set. *Educational and Psychological Measurement, 49*, 629-636.

Helm, S., Eggert, A., & Garnefeld, I. (2010). Modeling the impact of corporate reputation on

    customer satisfaction and loyalty using partial least squares. In V. E. Vinzi, W. W. Chin, J.

    Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods, and*

    applications (pp. 515-534), Berlin, Germany: Springer.

Hsu, S. H., Chen, W. H., & Hsieh, M. J. (2006). Robustness testing of PLS, LISREL, EQS and

    ANN-based SEM for measuring customer satisfaction. *Total Quality Management, 17*(3),

    355-371.

Hu, P. J. H., Clark, T. H. K., & Ma, W. W. (2003). Examining technology acceptance by school

    teachers: A longitudinal study. *Information & Management, 41*, 227-241.

Hubona, G. (2011). *Testing and assessing mediation in PLS path models* [Lecture notes].

    Chattanooga, TN: PLS-Seminars.

Human Resource Certification Institute. (2008, June 22). *HR Certification Institute unveils new*

    global branding campaign and reports 1,000[th] GPHR credential. Retrieved from

    http://www.hrci.org/aboutus/press/

Human Resource Certification Institute. (2011a). *HR certification statistics*. Retrieved from

    http://www.hrci.org/HRCertification.aspx?id=2147483768

Human Resource Certification Institute. (2011b). *PHR, SPHR and GPHR certification*

    handbook. Retrieved from http://www.hrci.org

Iacobucci, D. (2008). *Mediation analysis*. Los Angeles: Sage.

IBM. (2010). *SPSS command syntax to produce difference, standard error of the difference and*

    asymptotic 95% confidence interval on the difference between the zero order or bivariate

    correlation between v1 and v2 and the partial correlation between v1 and v2, partialling

    out v3. Methods from Olkin & Finn (1995). Retrieved from https://www-

    304.ibm.com/support/docview.wss?uid=swg21480367

Karweit, N., & Slavin, R. E. (1982). Time-on-task: Issues of timing, sampling, and definition.

    Journal of Educational Psychology, 74(6), 844-851.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New

    York: The Guilford Press.

Klomegah, R. Y. (2007). Predictors of academic performance of university students: An

    application of the goal efficacy model. *College Student Journal, 41*(2), 407-415.

Lane, J., Lane A. M., & Kyprianou, A. (2004). Self-efficacy, self-esteem and their impact on

    academic performance. *Social Behavior and Personality, 32*(3), 247-256.

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational*

    Behavior and Human Decision Processes, 50, 212-247.

Lee, C., & Bobko, P. (1992). Exploring the meaning and usefulness of measures of subjective

    goal difficulty. *Journal of Applied Social Psychology, 22*(18), 1417-1428.

Locke, E. A. (1991). The motivation sequence, the motivation hub, and the motivation core. *Organizational Behavior and Human Decision Processes, 50*, 288-299.

Locke, E. A., Frederick, E., Lee, C., & Bobko, P. (1984). Effect of self-efficacy, goals, and task strategies on task performance. *Journal of Applied Psychology, 69*(2) 241-251.

Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969-1980. *Psychological Bulletin, 90*(1), 125-152.

MacKinnon, D. P. (1994). Analysis of mediating variables in prevention and intervention research. In A. Cazares and L. A. Beatty (Eds.), *Scientific methods in prevention research*. NIDA Research Monograph 139, (DHHS Pub. No. 94-3631). Washington, DC: U.S. Government Printing Office.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York: Psychology Press.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods, 7*(1), 83-103.

McGill, T. J., & Klobas, J. E. (2009). A task-technology fit view of learning management system impact. *Computers & Education, 52*, 496-508.

McNatt, D. B. (2000). Ancient Pygmalion joins contemporary management: A meta-analysis of the result. *Journal of Applied Psychology, 85*(2), 314-322.

Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior, 9*, 596-606.

Mone, M. A. (1994). Comparative validity of two measures of self-efficacy in predicting academic goals and performance. *Educational and Psychological Measurement, 54*(2), 516-529.

Murphy, D., Campbell, C., & Garavan, T. N. (1999). The Pygmalion effect reconsidered: Its implications for education, training and workplace learning. *Journal of European Industrial Training, 23*(4/5), 238-250.

Nielsen, I. L., & Moore, K. A. (2003). Psychometric data on the mathematics self-efficacy scale. *Educational and Psychological Measurement, 63*(1), 128-138.

Oberst, J. E. (1995). *Seven principles student inventory: An indicator of success?* (Doctoral dissertation). Available from ProQuest Dissertations and Theses Database.

Olkin, I., & Finn, J. D. (1995). Correlations redux. *Quantitative Methods in Psychology, 118*(1), 155-164.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology, 21*(4), 325-344.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*(4), 426-443.

Phillips, J. M., & Gully, S. M. (1997). Role of goal orientation, ability, need for achievement, and locus of control in the self-efficacy and goal-setting process. *Journal of Applied Psychology, 82*(5), 792-802.

Ping, R. A. (2002). *Testing latent variable models with survey data: Appendix L – Average variance extracted.* Retrieved from http://www.wright.edu/~robert.ping/lv/toc1.htm

Pintrich, P. R., Smith, A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (Mslq). *Educational and Psychological Measurement, 53*, 801-813.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903.

Preacher, K. J. (2011). *Calculation for the Sobel test: An interactive calculation tool for mediation tests*. Retrieved from http://quantpsy.org/sobel/sobel.htm

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879-891.

Ringle, C.M., Wende, S., & Will, A. (2005). *SmartPLS 2.0 (M3)*. Retrieved from http://www.smartpls.de

Robertson, J. S., Grant, M. M., & Jackson, L. (2005). Is online instruction perceived as effective as campus instruction by graduate students in education? *The Internet and Higher Education, 8*, 73-86.

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science, 3*(6), 176-179.

Ryan, A. M., Ployhart, R. E., Greguras, G. J., & Schmit, M. J. (1998). Test preparation programs in selection contexts: Self-selection and program effectiveness. *Personnel Psychology, 51*(3), 599-621.

Rynes, S. L., Colbert, A. E., & Brown, K. G. (2002). HR professionals' beliefs about effective

    human resource practices: Correspondence between research and practice. *Human*

    Resource Management, 41(2), 149-174.

Schunk (1990). Goal setting and self-efficacy during self-regulated learning. *Educational*

    Psychologist, 25(1), 71-86.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory.

    Psychological Review, 76(2), 179-193.

Simonsson, M., Poelzer, H., & Zeng, L. (2000, April 24-28). *Teacher certification tests:*

    Variables that determine success for secondary pre-service teachers. Paper presented at

    the annual meeting of the American Educational Research Association, New Orleans, LA.

    Retrieved from http://www.eric.ed.gov

Sincoff, M. Z., & Owen, C. L. (2004). Content guidelines for an undergraduate human resources

    curriculum: Recommendations from human resources professionals. *Journal of Education*

    for Business, 80(2), 80-85.

Sosik, J. J., Kahai, S. S., & Piovoso, M. J. (2009). Silver bullet or voodoo statistics? A primer

    for using the partial least squares data analytic technique in group and organization

    research. *Group & Organization Management, 34*(1), 5-36.

Staples, D. S., & Seddon, P. (2004). Testing the technology-to-performance chain model.

    Journal of Organizational and End User Computing, 16(4), 17-36.

Tannenbaum, S. I., Mathieu, J. E., Salas, E., & Cannon-Bowers, J. A. (1991). Meeting trainees'

    expectations: The influence of training fulfillment on the development of commitment,

    self-efficacy, and motivation. *Journal of Applied Psychology, 76*(6), 759-769.

Temme, D., Kreis, H., & Hildebrandt, L. (2010). A comparison of current PLS path modeling software: Features, ease-of-use, and performance. In V. E. Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of partial least squares: Concepts, methods, and applications* (pp. 737-756), Berlin, Germany: Springer.

Titard, P. L., & Russell, K. A. (1989). Factors affecting CPA examination success. *Accounting Horizons, 3*(3), 53-59.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186-204.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables. *Journal of Applied Psychology, 76*(6), 759-769.

Wiley, C. (1995). Reexamining professional certification in human resource management. *Human Resource Management, 34*(2), 269-289.

Wiley, C. (1999). A comparative analysis of certification in human resource management. *The International Journal of Human Resource Management, 10*(4), 737-762.

Wood, R. E., & Locke, E. A. (1987). The relation of self-efficacy and grade goals to academic performance. *Educational and Psychological Measurement, 87*(4), 1013-1024.

Yi, M. Y., & Hwang, Y. (2003). Predicting the use of Web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International Journal of Human-Computer Studies, 59*, 431-449.

Yu, T. K., & Yu, T. Y. (2010). Modeling the factors that affect individuals' utilisation of online learning systems: An empirical study combining the task technology fit model with the theory of planned behavior. *British Journal of Educational Technology, 41*(6), 1003-1017.